Report 86Y061

IMAGE METRICS

A. B. Lucero, G. B. Silverman, J. W. Bair, and R. R. Ramroth
Northrop Corporation
Electro-Mechanical Division
500 East Orangethorpe Avenue
Anaheim, CA 92801

1 May 1986

Interim Technical Report for Period October 1985 through April 1986

Prepared for
NIGHT VISION & ELECTRO-OPTICAL CENTER
Fort Belvoir, VA 22060

Approved By:

Guner S. Robinson

Guner S. Robinson, Director
Advanced Technology Laboratory

SUMMARY

Background

Most Automatic Target Recognizer (ATR) development and testing has been conducted in the community without a general means for objectively and quantitatively evaluating performance. It is clear that the performance of any ATR is dependent upon the quality of the input signal. For example, as the noise or clutter in an image becomes increasingly severe, the ATR performance will correspondingly degrade. Therefore it is important to have an objective and quantitative measure of image signal quality by which ATR performance may be gauged.

Objectives

The objectives of this work are to provide and validate a means for quantitative measurement of image signal quality, for use in evaluation of ATR's.

In this work we address the ATR function of detection, as performed by a front-end region of interest operator or interest point locator. We also treat both targets and clutter in order to address both true detection rates and false alarm rates.

This work is to provide a limited analysis of existing image metrics. Target-to-interference ratio (TIR) is the principal metric. Alternate forms of TIR are also proposed and analyzed.

In addition, a limited validation exercise is performed for the TIR metrics, as well as for metrics developed under an Air Force contract called Criteria for Target Recognizer Evaluation (CTRE). The validation exercise consists of determining the degree of correlation between image metrics and the Haar region of interest operator. Two versions of the region of interest operator are used; both are based on a Haar detector that is used in the Army Automatic Target Cuer (ATC) ATR. The emphasis in the validation exercise is to uncover any serious deficiencies in the image metrics and rectify those found.

Technical Problems

In analyzing image metrics, the technical problems we faced included: (1) establishing design criteria for the operators used to compute image metrics; (2) establishing a fully automated algorithm for raw image metrics; (3) establishing a procedure for taking into account the variable size target image that occurs with change in target orientation, range-to-target, and sensor depression angle; (4) taking account of the effect of nearby objects on the TIR values of one another; and (5) taking into account the effect of a realistic sensor on TIR estimates.

In validating image metrics, the technical problems that we faced included: (1) rendering the Haar detector and the image metrics to a common base so that meaningful correlation analyses could be performed; (2) choosing an appropriate correlation analysis method that would not require assumptions regarding a mathematical functional form that relates detector outputs to image metrics; (3) treating two points of view of the correlation analysis–all possible objects prior to Haar detection or, alternatively, only those objects that are actually detected by Haar; (4) selecting the appropriate measurements for correlation analysis–namely individual (raw) measurements or aggregate (smoothed) measurements; and (5) providing a useful validation result within available computer resources.

General Methodology

The methodology employed to solve these technical problems, in the order listed, involved: (1) development of a Haar figure of merit that is gain and offset invariant, as are the image metrics; (2) application of Kendall's nonparametric test for rank order correlation between detector outputs and image metrics; (3) use of a low threshold on the image metrics for correlation analysis of objects prior to Haar detection and, alternatively, use of a low threshold on Haar detections for correlation analysis of objects, after Haar detection; (4) individual (raw) measurements consisted of each measurement pair of (Haar figure of merit, image metric value) associated with each individual object, whereas aggregate (smoothed) measurements consisted of each parametric pair of (average Haar figure of merit, threshold value on image metric) associated with each subset of objects whose image metric value exceeded the indicated threshold; (5) given various approximations (regarding image

metrics and spatial registration of image metric reports with detection reports), the results obtained were reasonable and served their intended purpose.

Technical Results

The technical results for analysis of image metrics include the following: an algorithm was established for specifying the size and shape characteristics of the operators used to extract information for computation of image metrics, taking realistic sensor characteristics into account. Mathematically, TIR was found to exhibit an asymmetry with respect to errors produced by a mismatch with actual target size. A viable post-processing algorithm was developed that is fully automatic, takes into account the influence of nearby interfering objects, and produces the desired image metric report for objects of interest (targets or non-targets).

The technical results for validation of image metrics include the following: Raw and post-processed image metric maps were produced for 50 images, for two (plus) metrics. Raw and post-processed Haar detection results were also produced for the same 50 images, for two versions of the Haar detector. Sixty-four separate data base correlations were made with image metrics versus Haar detection reports. Correlation values and confidence levels were obtained for each.

Important Findings and Conclusions

In general TIR correlated somewhat better with Haar than did CTRE. (The alternate forms of TIR did not correlate well.) For correlations against targets only, the average TIR correlation was 0.94; and the average CTRE correlation was 0.92 (where 1.00 is perfect correlation). For correlations against non-targets only, the average TIR correlation was 0.79; and the average CTRE correlation was 0.72. On a single–experiment basis, the differences between TIR and CTRE correlation values were not significant, except for one case. If one had to make a choice between TIR and CTRE, TIR should be accepted, primarily on the basis of lesser computational complexity.

As is evident from these correlation values, a useful relationship has been demonstrated for both TIR and CTRE as both predictors and explainers of Haar

detection performance. The aggregate results do not appear to be very sensitive to the precise shape and size of the operator masks used.

Based on these experiments, as well as prior work in this area, we expect that the major objective of this effort has been achieved; namely, a means has been provided and validated for quantitative measurement of signal quality (for both targets and clutter) for use in ATR evaluation.

PREFACE

This study was initiated by the Defense Advanced Research Projects Agency (DARPA) as part of the Passive Autonomous Infrared Technology Program (PAIRSTECH). It is directed by the U.S. Army Night Vision and Electro-Optics Center (NVEOC), Fort Belvoir, Virginia 22060. The study is being conducted by the Northrop Corporation, Electro-Mechanical Division, 500 East Orangethorpe Avenue, Anaheim, California 92801, under Contract No. DAAL02-85-C-0165. Mr. Antonio B. Lucero, Principal Systems Engineer, is the Project Manager for Northrop Corporation. The work reported herein was started in October 1985 and was completed in April 1986. This report was submitted to NVEOC by Mr. Lucero in May 1986.

ACKNOWLEDGEMENTS

Many thanks are due to several people who contributed to the review of this report, especially to Phyllis Lewis, editor, for her helpful suggestions. Special thanks are due to Mike Cook for his conscientious work in coordination and production of much of the photographic artwork. Most of all, we are extremely appreciative of the dedication and efficiency demonstrated in the preparation of this report by our exemplary secretary Linda Adkins.

TABLE OF CONTENTS

Section		Page
1	INTRODUCTION	11
2	TARGET-TO-INTERFERENCE RATIO (TIR) FORMULATIONS AND CONCEPTS	13
2.1	TIR History	13
2.2	Radar Background	13
2.3	Extension to Imaging Electro-Optical Systems	15
2.4	Alternate Forms to TIR	17
3	TIR APPLICATION AND PROCEDURES	20
3.1	Previous Application	20
3.2	Current Application	22
3.3	Adopted Procedure	24
4	BRIEF SUMMARY OF CTRE METRICS	28
5	IMAGE METRIC VALIDATION	31
5.1	Approach	32
5.1.1	Basic Approach	32
5.1.2	Imagery Selection	33
5.1.3	Haar Detection Reports	37
5.1.4	Image Metric Reports	38
5.1.5	Registration of Haar Detection Reports with Image Metric Reports	41
5.1.6	Correlation Analysis Method	43
5.2	Results	45
6	CONCLUSIONS AND RECOMMENDATIONS	59
	REFERENCES	61

APPENDICES

Section		Page
Appendix A	Estimate of σ_B for TIR When Limited by Quantization Step Size	62
Appendix B	TIR as a Test for Zero Contrast	64
Appendix C	Sensitivity of TIR to Target Size Estimation Errors	66
Appendix D	Robust TIR	72
Appendix E	ATC Firmware Version of the HAAR Detector	73
Appendix F	HAAR Figure of Merit	81
Appendix G	Relationship Between Kendall's τ and the Sample Correlation Coefficient	83
Appendix H	Numerical Example of Kendall's τ and Associated Parameters	85

LIST OF ILLUSTRATIONS

Figure		Page
1	Spatial Filter	16
2	Effect of a Point Spread Function	24
3	Original ERIM - Supplied Image	26
4	Ground Truth Target Regions	26
5	TIR Map (Exponential Scaling)	27
6	Post-Processed TIR Map	27
7	Procedure for Determining Regression Model for Prediction of Detection Performance as a Function of Image Quality	29
8	Target and Background Sampling	30
9	Typical Characteristic Curves	34
10	Example of Raw Measurements	46
11	Example of Smoothed Measurements Obtained from Figure 10	46
12	Simplified Pseudo Code Describing Process	47
13	TIR New Map (Exponential Scaling)	48
14	Post-Processed TIR New Map	48
15	CONTRAST Map	49
16	Post-Processed CONTRAST Map	49
17	CTRE Map	50
18	Post-Processed CTRE Map	50
19	Unnormalized HAAR Filter Map	51
20	Detections Using Low Percentile Threshold on Unnormalized HAAR Map	51
21	HAAR Figure of Merit (HFOM) Map	52
22	Detections Using Low Threshold on HFOM Map	52
E-1	HAAR Detector	74
E-2	ATC HAAR Mask Implementation	75
E-3	Typical Histogram of HAAR Pixel Output	77
E-4	Coloring Flow Diagram	79
E-5	Coloring Logic Requirements	80

LIST OF TABLES

Table		Page
1	Large Target Validation Data Base	35
2	Small Target Validation Data Base	36
3	Detection Reports for ERIM Frame #2007028	39
4	Relative Differences in Target Areas with Respect to Reference Targets	41
5	Image Metric Reports for ERIM Frame #2007028	42
6	A Posteriori Correlation of CONTRAST with ATC HAAR FOM	53
7	A Posteriori Correlation of CONTRAST with HAAR FOM	53
8	A Posteriori Correlation of TIR_{NEW} with ATC HAAR FOM	54
9	A Posteriori Correlation of TIR_{NEW} with HAAR FOM	54
10	A Posteriori Correlation of TIR with ATC HAAR FOM	55
11	A Posteriori Correlation of TIR with HAAR FOM	55
12	A Posteriori Correlation of CTRE with ATC HAAR FOM	56
13	A Posteriori Correlation of CTRE with HAAR FOM	56
14	A Priori Correlation of TIR with HAAR FOM	57
15	A Priori Correlation of CTRE with HAAR FOM	57
H-1	Integral Measure Data for TIR on Small Target Image Set	86

1. INTRODUCTION

A great deal of automatic target recognizer (ATR) development and testing have been conducted in the community without a general means for objectively and quantitatively evaluating performance. It is clear that the performance of any ATR is dependent upon the quality of the input signal. For example, as the noise or clutter in an image becomes increasingly severe, the ATR performance will correspondingly degrade. Therefore it is important to have an objective and quantitative measure of image signal quality by which ATR performance may be gauged.

The objectives of this work are to provide and validate a means for quantitative measurement of image signal quality, for use in evaluation of ATR's.

In this work we address the ATR function of detection, as performed by a front-end region of interest operator or bulk filter. We also treat both targets and clutter in order to address both true detection rates and false alarm rates.

In this report, we investigate image signal quality metric formulations and concepts, as well as procedures for applying them to the current application. We also carry out an image metric validation exercise and show quantitative correlation results against an actual ATR detection algorithm designed for forward-looking scenarios.

The background for the present study is documented in References 1 through 4. In Reference 1, preliminary definitions of image metrics were provided. In Reference 2, calibration runs were made for targets with additive noise. In Reference 3, correlations were made with an actual ATR designed for down-looking scenarios. In Reference 4, correlations were made with human detection performance.

In this report, we extend the metric definitions and analyze each metric in greater depth, bringing general ATR knowledge to bear. Extensive nonparametric correlation analyses are performed using Kendall's Test to demonstrate the effectiveness of the image metrics.

In Section 2 we discuss the work performed in the formulations and concepts underlying the baseline image metric. In Section 3 we discuss applications and

procedures for obtaining image signal quality metric values for targets and non-targets. In Section 4 we provide a brief summary of the metrics used in Reference 4.

In Section 5 we provide a detailed description of the validation exercises conducted against an actual ATR, as well as a summary of the results obtained.

In Section 6 we present our conclusions and recommendations.

2. TARGET-TO-INTERFERENCE RATIO (TIR) THEORY AND CONCEPTS

TIR is a measure of image signal quality by which the performance of an ATR may be objectively evaluated. In this section we briefly describe the history of TIR, its conceptual genesis in radar signal quality estimation, its extension to imaging systems, and alternate forms to TIR.

2.1 TIR History

Northrop Research and Technology Center (NRTC) first developed and applied the TIR metric to image signal quality measurement and ATR evaluation in 1980 (References 1 and 2). From the onset, TIR was intended as a simple, straightforward indicator of signal quality, rather than an elaborate algorithm rivaling the complexity of an ATR. It was used successfully to evaluate ATR performance on an IR&D project Northrop Advanced Tactical Seeker (NATS); and later on a DARPA/MICOM contract Advanced Sensor Signal Processor (ASSP), DAAH01-82-C-A106 (Reference 3).

Then, in 1983, Northrop Electro-Mechanical Division again applied TIR (and other metrics) to image signal quality measurement and correlated it with target recognition performance of humans. This was done on an Air Force contract Criteria for Target Recognizer Evaluation (CTRE), F33615-83-C-1094 (Reference 4).

2.2 Radar Background

The concept underlying TIR stems from the signal-to-noise ratio (SNR) used in radar to characterize signal quality.

$$SNR = \frac{E(P_S)}{S(P_N)}; \qquad (2\text{-}1)$$

where $E(P_S)$ = expected value of received signal (target) power, and

$S(P_N)$ = standard deviation of additive noise power.

In this expression the assumptions are that

$$E(P_N) = 0,$$

$$S(P_N) = S(P_S) \text{ or } S(P_N) >> S(P_S),$$

and the samples are independent.

A more general form would be:

$$SNR = \frac{E(P_S) - E(P_N)}{S(P_N)} \qquad (2\text{-}2)$$

for non-zero mean noise and $E(P_S) > E(P_N)$.

The SNR is used to indicate the level of difficulty in extracting the signal from the noise; in particular, the probability of detecting the signal is a monotonically increasing function of SNR:

$$\Pr(\text{Signal Detection}) \{ \text{ SNR} \qquad (2\text{-}3)$$

where the symbol "{ " is used to denote a positive monotonic relationship.

In the presence of clutter (or even another close-by target), an important influence on signal quality is the interference caused by such objects. When the influence of clutter dominates over the influence of noise, an analogous expression has been used in radar, namely signal-to-clutter ratio (SCR).

$$SCR = \frac{E(P_S) - E(P_C)}{S(P_C)}; \qquad (2\text{-}4)$$

where $E(P_C)$ = expected value of the clutter power,

$S(P_C)$ = standard deviation of the clutter power,

and $E(P_S) > E(P_C)$.

When both clutter and noise interfere with signal detection, the combination of the two is termed "interference." For additive noise, the composite power of the interference, P_I, is modeled as the noncoherent sum of their powers

$$P_I = P_C + P_N.$$

So analogously, the signal-to-interference ratio (SIR) is given by

$$SIR \approx \frac{E(P_S) - E(P_I)}{S(P_I)}, \qquad (2\text{-}5)$$

where $E(P_S) > E(P_I)$

For noise only, SIR reduces to SNR; and for clutter only SIR reduces to SCR.

2.3 Extension to Imaging Electro-Optical Systems

In both FLIR and television systems, the output image is noncoherent and the display's brightness is a monotonic function of the incident power on the collecting optics. However, unlike the radar case outlined above, the signal is usually not represented by a single point measurement (i.e., a single pixel). Instead it is represented by a collection of more or less connected pixels corresponding to a target image.

If we know a priori the target size to be expected, we can improve the signal quality by bandpass filtering the spatial frequencies corresponding to a target (analogous to a radio receiver tuning to an expected portion of the radio frequency band).

So, for example, if a target is expected to have a length ℓ and a height h, a spatial filter or mask can be defined which matches the expected target dimensions; and, under the hypothesis that a target is present, an associated spatial filter or mask can be defined which samples the clutter and/or noise in a neighborhood around the hypothesized target. The combination of these masks is shown in Figure 1. A guard region may be defined between the target mask and the background mask to reduce the chance of mixing samples of target and background, which would tend to increase the error in parameter estimation for each population.

Since, in the case of electro-optical (E-O) imagery, the spatial filtering is matched to expected target dimensions, this lead us to adopt the usage of "target" in place of

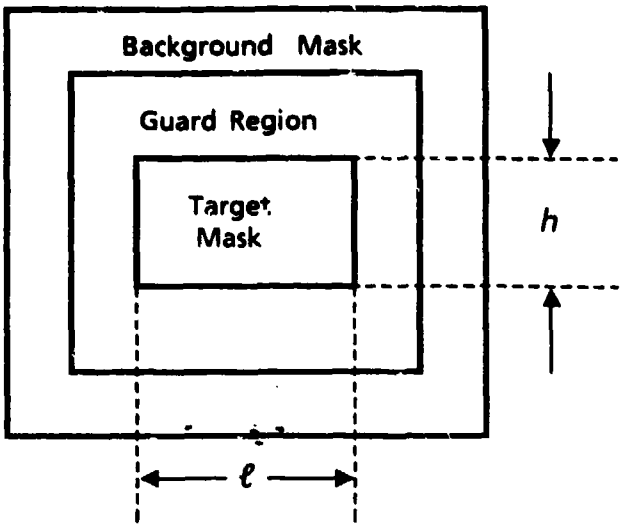


Figure 1. Spatial Filter

"signal." Thus, the target-to-interference ratio (TIR) for E-O imagery is defined as

$$TIR \equiv \frac{|\mu_T - \mu_B|}{\sigma_B}, \tag{2-6}$$

where $\mu_T \equiv$ the average value of the pixel intensities (or grey levels) in the target mask,

$\mu_B \equiv$ the average value of the pixel intensities in the background mask,

and $\sigma_B \equiv$ the sample standard deviation of the pixel intensities (clutter + noise) in the background mask.

This expression can also be thought of in terms of contrast relative to the mean background level, which could be expressed as

$$TIR \equiv \frac{|\,average\ target\ contrast\,|}{background\ standard\ deviation} \tag{2-7}$$

In Eq. (2-6), the assumption is that the standard deviation of the target pixels, σ_T, does not affect target detectability. (This assumption will be addressed at length in later sections of this report). TIR is basically the z-score of the average target intensity relative to the background population; also $(TIR)^2$ is the Mahalanobis distance of the average target intensity from the background population.

2.4 Alternate Forms to TIR

There are several shortcomings with Eq. (2-6) and the assumptions underlying it. On the other hand, in defining an image signal quality measure, we originally set out to establish a straightforward, simple indicator, rather than an elaborate algorithm. So, as long as the shortcomings are understood, and are not of a catastrophic nature, the TIR metric may still serve the intended purpose. The following discussion describes the TIR shortcomings.

First, as the standard deviation of the background, σ_B, approaches zero, TIR approaches infinity. It is possible that with a uniform background and $\sigma_{noise} <$ quantization step size, σ_B could be computed as zero. If such an event were to occur, the logical thing to do would be to substitute a finite value for σ_B, say σ'_B, that is consistent with the σ_B estimate of zero and the quantization step size, while preventing TIR from becoming excessively large.

Based on an analysis in Appendix A, "Estimate of σ_B for TIR When Limited by Quantization Step Size," a reasonable value for σ'_B is 0.24 of a grey level. It represents an upper bound for six cases covering different population sizes, uniform and normal distributions, for 90% and 99% confidence intervals.

Another shortcoming for which there is no simple remedy is the case where the numerator of Eq. (2-6) is zero and, therefore, TIR is zero, even though one can clearly discern the target against its background. Other forms such as the Bhattacharyya measure have been considered for elimination of this shortcoming; however, due to time limitations, this alternative form was not investigated.

Also if either the background or target pixel intensities do not come from a normal distribution, say they are multimodal, then the parametric form of Eq. (2-6) may be misleading. In such a case, nonparametric measures are more appropriate such as the general Bhattacharyya measure or a Kolmogorov-Smirnov (K-S) type of test.

These alternatives are not simple changes to make. However, some of the CTRE metrics are already based on K-S tests; consequently, it would be interesting to compare CTRE results with TIR results for such cases.

Spatial correlation of pixels in FLIR imagery is the rule rather than the exception. However, estimates of μ_T, μ_B, and σ_B are made for TIR, based upon the assumption of independent samples. At the present time, we are not sure of the full effect this has on TIR estimates. One effect, however that this would surely have is on the confidence interval about TIR estimates. This would also tie in with understanding the effect that target size has on TIR. If, for example, the target consists of only two or three independent pixels, then the confidence intervals will be large, and, therefore, the detectability will be uncertain, even for a high-contrast target.

Another potential shortcoming of TIR is that the standard deviation of the target pixels, σ_T, is not included in Eq. (2-6). Under the assumption of normal distributions, perhaps the appropriate formulation for TIR which includes σ_T would be

$$TIR_{new} = \frac{|\mu_T - \mu_B|}{\sqrt{\sigma_T^2 + \sigma_B^2}} . \tag{2-8}$$

From a theoretical point of view this has more appeal than Eq. (2-6), since it forms a valid test statistic for testing the hypothesis that there is no significant contrast between the populations of target pixels and background pixels. Extending this idea further, another alternate form is suggested, namely, CONTRAST.

$$CONTRAST = |\mu_T - \mu_B| - \left(\frac{\sigma_T^2}{n_T} + \frac{\sigma_B^2}{n_B}\right)^{\frac{1}{2}} Z_\alpha , \tag{2-9}$$

where n_T = number of target pixels,

n_B = number of background pixels,

Z_α = the value of standard normal deviates which correspond to the α confidence level. (See Appendix B for more detail.)

CONTRAST represents a conservative (lower bound) estimate of the true contrast. The way in which Z_α is used to determine this lower bound is shown with an

illustrative example. Suppose, as we did in this study, that the lower bound of the 95% confidence interval were desired. Then Z_α is the value of the normal deviate (zero mean and unit variance) such that 97.5% of the standard normal distribution is larger than Z_α. Using the table of standard normal deviates, one finds a table entry of 0.475 (95.0/2). This corresponds to a value for Z_α of 1.96. This would be substituted in Eq. (2-9) to determine the value of CONTRAST.

3. TIR APPLICATION AND PROCEDURES

In the application of TIR to actual imagery, there are a number of practical considerations:

a. How should the triple-gate TIR mask be designed?

b. How should TIR be applied?

c. How will variable target size and variable target aspect angle within a frame affect TIR?

d. How will nearby objects affect local TIR estimates?

e. How will the sensor point spread function affect TIR?

In this section we will address each of these questions. First we will review how these were addressed in work done prior to this study. Then we will describe important differences regarding image metrics for our current application versus the prior work. Finally, taking all these considerations into account, the procedure for computing TIR for this study is described.

3.1 Previous Application

Our work on TIR in the early 1980's established a standard approach for each of the above questions, except for questions d and e. Our approach on this project builds upon that experience. What had been done previously applied to a top-view scenario for a limited altitude interval in which the target size would vary by no more than $\pm 15\%$.

For that scenario, the size of the innermost gate was chosen to correspond to the expected target size; and the guard region was chosen to exceed the maximum target size for any target rotation. The nominal target used for these estimates was a tank since it was the most important target type. A set of eight masks, at rotation intervals of $22\frac{1}{2}$ deg, formed a filter bank. Each filter in the bank was applied as a convolutional operator over the whole frame. At each pixel location, the maximum

output value across all eight filters was selected as the TIR value reported for that pixel.

This approach was susceptible to contaminated estimates of μ_T. This could occur whenever the target was smaller than the innermost gate or whenever the target rotation angle was other than a multiple of $22\frac{1}{2}$ deg. μ_B and σ_B, however, were never contaminated, unless an extremely long truck happened to be the target. This mask design never appeared to cause a very significant problem in obtaining a reasonable estimate for TIR.

An interesting phenomenon that is worth mentioning at this point is that very often the peak TIR value for a target would not occur when the TIR mask was centered on a target, but would be offset (within the target interior though). The cause of this phenomenon often was nearby interfering objects which produced a larger value of σ_B (and, therefore smaller TIR) when the TIR mask was centered on the target than when it was off center.

There were, however, two types of problems in which our TIR formulation and/or mask design did not produce consistent results with respect to actual ATR detection outputs. These were: (1) targets parked adjacent to high contrast extended clutter, and (2) very–low–contrast targets or very–high–contrast targets. In case (1), apparently the clutter caused σ_B to be large, and this in turn caused TIR to become exceedingly low. In case (2), as compared with ATR figures of merit, low–contrast targets in competing clutter yielded extremely low TIR values; and high–contrast targets in benign clutter yielded extremely high TIR values. This lack of correspondence at extreme ends of signal quality was attributed to TIR being very simple. As compared with the ATR, TIR was not as good at detecting weak targets, and its quality measure was easier to satisfy with only moderately strong targets.

There was also a procedural problem with mapping pixel-level TIR outputs to higher object-level TIR values. The procedure that was followed was to perform spatial clustering (or coloring); then, by visual inspection, certain clusters were eliminated based on: (1) sizes that were significantly larger than the typical target response or were single–pixel responses; and (2) aspect ratios that were much larger than those of typical target responses. By probing pixel TIR values in a neighborhood, an estimate of the peak TIR value was obtained within each surviving cluster and was then reported as the object TIR value. In some cases, where a target cluster merged

with the cluster of adjacent clutter, the estimated peak TIR value in the vicinity of the target was reported as the target TIR.

3.2 Current Application

Turning to the application at hand, we have a forward–looking scenario in which the range-to-target within a single frame can vary considerably, depending on the depression angle of the sensor. And, even at a specific range-to-target, the apparent target horizontal extent can change by 100% as the target heading changes from end-view to broadside. The problem is further complicated by the projection from three dimensions to two dimensions, with considerable target variability occurring in the vertical direction of the display. For this scenario the characteristic target dimensions are projected length ℓ (horizontal direction) and projected height h (vertical direction).

In order to design the TIR mask to best match the target size and shape, while accommodating apparent target variability and yet minimizing the sensitivities in TIR estimates, it was necessary to conduct an analysis of the effect of contamination of the target region by background pixels, and vice versa. This analysis is described in Appendix C, "Sensitivity of TIR to Target Size Estimation Errors." The conclusion of this analysis was that the impact of contamination is asymmetrical for the original TIR form of Eq. (2-6).

For example, suppose the ideal TIR is 3.16. Then for the case where 10% of the target region contains background pixels, TIR would be underestimated by 10%. On the other hand, for the case that 10% of the background region contains target pixels, TIR would be underestimated by 34%. In general, in either case, TIR would be underestimated. From this analysis one would conclude that it is more costly to contaminate the background region than the target region.

Reinforcement of this conclusion arises from consideration of the frequently occurring imaging infrared phenomenon, in which the range of intensities of target pixels is often quite broad and contains the range of local background intensities. Thus, contamination of the background region by target pixels is much more likely to significantly influence the estimates for μ_B and σ_B, than would contamination of the target region by background pixels influence the estimate of μ_T.

Consideration was also given to making the estimates of μ_B, μ_T, and σ_B more robust, i.e., less sensitive to contamination. (See Appendix D, "Robust TIR.") However, it was concluded that it would not be a suitable approach because it would eliminate or reduce the true influence of a target "hot spot" or a bimodal background on ATR target detectability. These are obviously important phenomena that should not be suppressed. Therefore, a robust TIR was rejected as a possible alternate approach.

The premise of designing the TIR mask that best matches the target size and shape was then re-examined. As previously described, the TIR mask definition included a guard region to protect against, or diminish, the effects of mixing target (/background) pixels into the background(/target) region. This approach to solving the problem of population mixing makes sense when the target size and shape are known a priori to a fair degree of precision (as was the case of the top-view, constant-range scenario of NATS/ASSP). In such cases, rejecting a ring of one or two pixels between the target region and background region would be expected to improve the TIR estimate.

Unfortunately, in our current application there is a great deal of variation of target size and shape and a great deal of uncertainty about the size and shape expected for any given target; so it is unlikely that an effective guard ring can be modeled.

The original modeling done for the TIR mask also considered only geometric factors. The effect of the sensor point spread function was therefore not taken into account. From the TIR point of view, the sensor point spread function has the effect of smearing an image, as compared with an ideal geometric representation. This smearing causes adjacent pixels to share intensities. As a result, pixels within the target's geometric extent and pixels outside of the target's geometric extent share their intensities. The closer they are, the more sharing will take place. Figure 2 illustrates the effect of a point spread function on an ideal rectangular waveform.

In general, the wider the point spread function, the more difficult it will be to discriminate targets from their backgrounds.

If we now re-examine the role of a guard region, it is no longer clear what should be considered exclusively target region and what should be considered exclusively background region. And even if we utilized a large guard region which excludes

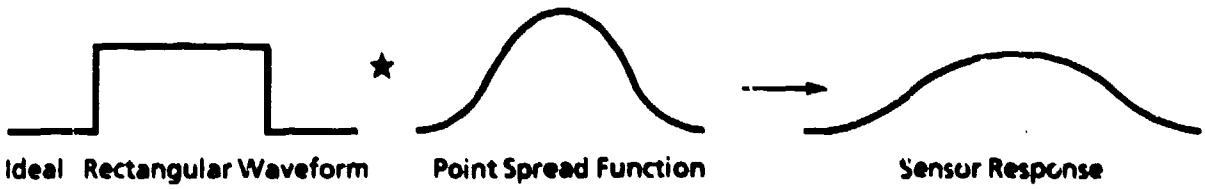


Figure 2. Effect of a Point Spread Function

most of the target/background mixing, such a guard region would also hide the effect on target detectability of the width of the point spread function.

Taking these factors into account, it was decided that we would compute TIR without using a guard ring; we would use a bank of masks that have only two regions: target and background. The mask sizes and shapes would be designed to include a target region at least as large as the largest possible target size. This approach allows for at least one mask that attenuates the effect of target pixels contaminating the background region (the most serious type of error in TIR estimation). The maximum TIR over all the masks is then reported as the TIR value for the pixel corresponding to the center of the masks.

3.3 Adopted Procedure

Taking the foregoing considerations into account, the procedure for computing TIR for the current scenario is as follows: A bank of nine parametrically range-scaled filters is defined, one for each of nine target headings, from head-on aspect through broadside aspect. At a specific range and at each aspect angle, the rectangular target region is designed to include the largest target projection, and by complementarity the background region is designed to exclude the largest target projection. All nine masks are applied at each pixel location, since it is not known a priori what the target headings are going to be. The largest TIR computed among the nine is then selected as the reported TIR value.

As the estimated range-to-target changes, the filters are scaled to reflect the expected target size, within a specified range interval of interest. The result of applying the TIR bank of filters is a map of TIR values defined at each pixel.

For purposes of illustration, examples are shown in Figures 3 through 5 of an original grey-level image, an image showing ground-truth target regions, and the corresponding TIR map. (Although this TIR map was obtained using only a single TIR filter, rather than a bank of filters, its qualitative appearance is identical to the filter bank output).

The next step is post-processing of the TIR map to determine the maximum TIR value within the spatial extent of each target-sized object. We have developed an automated method for extracting the maximum object-level TIR value. This method utilizes the "max operator."

Given a maximum projected target size of height h and length ℓ, the TIR map is post-processed first by a $1 \times \ell$ nonlinear convolutional mask that outputs the maximum TIR value contained within the 1 row by ℓ columns. This results in a new map which is then convolved with another max operator of dimension $h \times 1$. It outputs the maximum value within its h rows by 1 column.

The final result is a max TIR map composed of rectangular plateaus of constant (max) TIR values. Figure 6 shows results from post-processing Figure 5. Each plateau corresponds to a local maximum TIR value. The center of each complete plateau is the location of the pixel in the original TIR map with the local maximum TIR, and the TIR value of the plateau is, in fact, the value of the local maximum TIR.

This post-processing algorithm completely disassociates target-size objects with larger TIR values from objects with smaller TIR values. Since targets usually have a larger TIR value than nearby clutter, this is a desirable attribute. If, on the other hand, an object of interest has a smaller TIR than another, nearby object, then its plateau may be partially masked by the plateau of the other object. (In this case the center of the lower TIR plateau will not be the location of its source.) If a plateau does not contain its own source point, then there is an object within half the target dimension that has a larger TIR. In this case the partial plateau is rejected. If, however, a plateau does contain its own source point, it is retained.

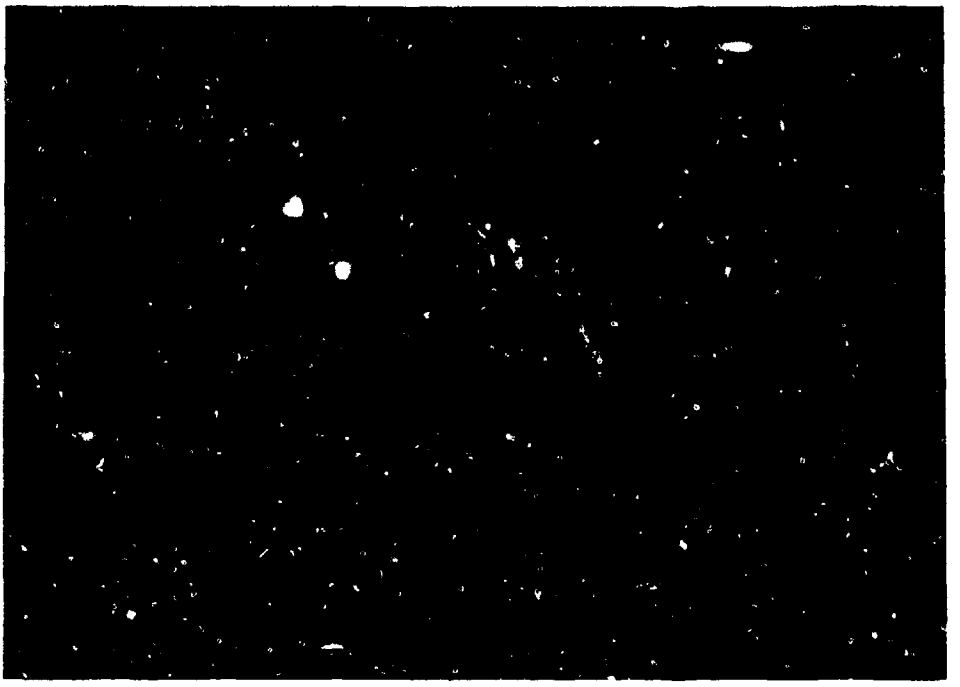
The efficiency and objectivity of this post-processing algorithm is a significant extension of the previous methodology for estimating image signal quality.

BS 731 A789-1

Figure 3. ORIGINAL ERIM - SUPPLIED IMAGE

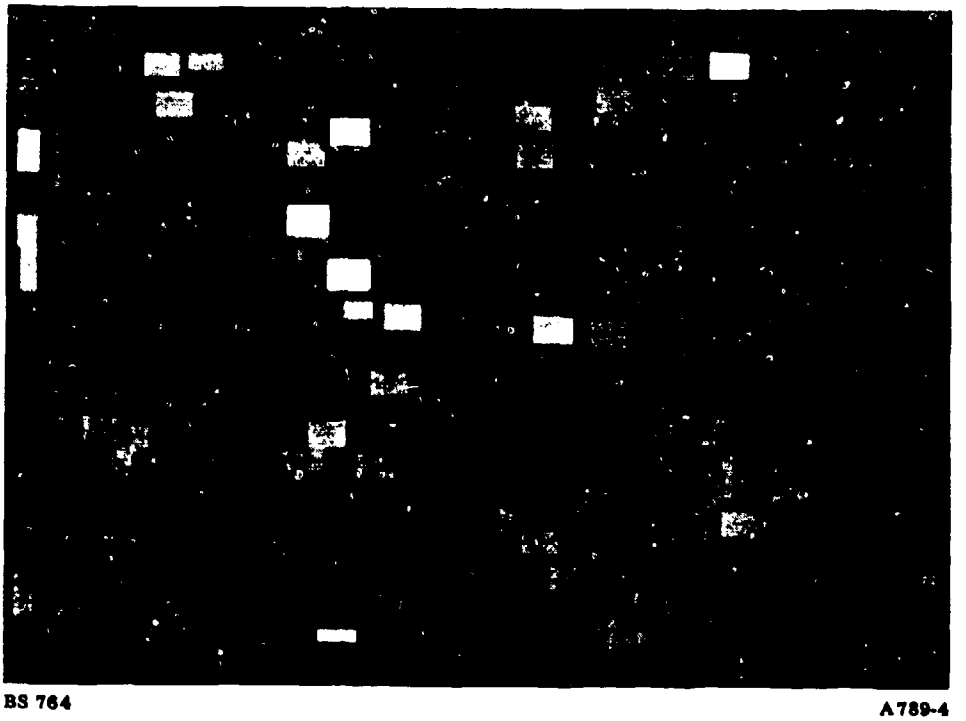
BS 771 A789-2

Figure 4. GROUND TRUTH TARGET REGIONS



BS 773 A789-3

Figure 5. TIR MAP (EXPONENTIAL SCALING)



BS 764 A789-4

Figure 6. POST-PROCESSED TIR MAP

4. BRIEF SUMMARY OF CTRE METRICS

In addition to TIR and TIR-related metrics, we also considered metrics developed under the CTRE project. (For full details see Reference 4.)

The CTRE technical approach consisted of (1) proposing an initial set of candidate metrics based on established hypothesis tests computed on an image window and (2) performing regression analysis of human subjects' target detection scores versus alternative subsets of candidate metrics. By demonstrating the ability of metrics to predict the probability with which targets are detected by humans, it was expected that they would be useful for predicting ATR performance. In the end, four of the original eight metrics (based on contrast or edge-gradient magnitude) were found to account for 91 percent of the variation in the human subjects' detection performance.

Initially, regression analysis was performed against individual local image metrics. Many of the metrics were found to have significant regression, but no individual one was found to be an adequate predictor. Later research focused on multivariate regression analysis and variable subset selection procedures.

The multivariate regression analysis consisted of: (1) evaluating alternative subsets of the metrics; (2) filtering the experimental data to suppress outliers; and (3) performing weighted least squares regression on the filtered data to produce the final model.

The fitting of human performance scores with a simple linear regression model posed severe technical problems with respect to outliers. However, instead of judging which data points were outliers and rejecting them from the regression analysis prior to demonstrating any significant regression, a method of soft rejection based on "robust" regression techniques was employed. This was followed by the application of weighted least squares.

A diagram illustrating this procedure is shown in Figure 7.

The regression model for probability of detection, P_D, that resulted from this procedure is given by

$$P_D = 0.3 + 1.08m_1 + 0.072m_2 + 0.794m_3 - 0.807m_4, \qquad (4\text{-}1)$$

where m_1 = Normalized Chi^2 test on grey shades (measures the frequency and degree to which target and background grey shade histograms are dissimilar),

m_2 = Normalized Kolmogorov-Smirnov (K-S) test on edge gradients (measures the probability that the target's gradient distribution tends to have stronger gradients than that of the background),

m_3 = Normalized (TIR)^2 (measures the squared ratio of the contrast between target and background to the standard deviation of the background),

and m_4 = Normalized K-S test on grey shades (measures the probability that the target's grey shades tend to be lighter than those of the background).

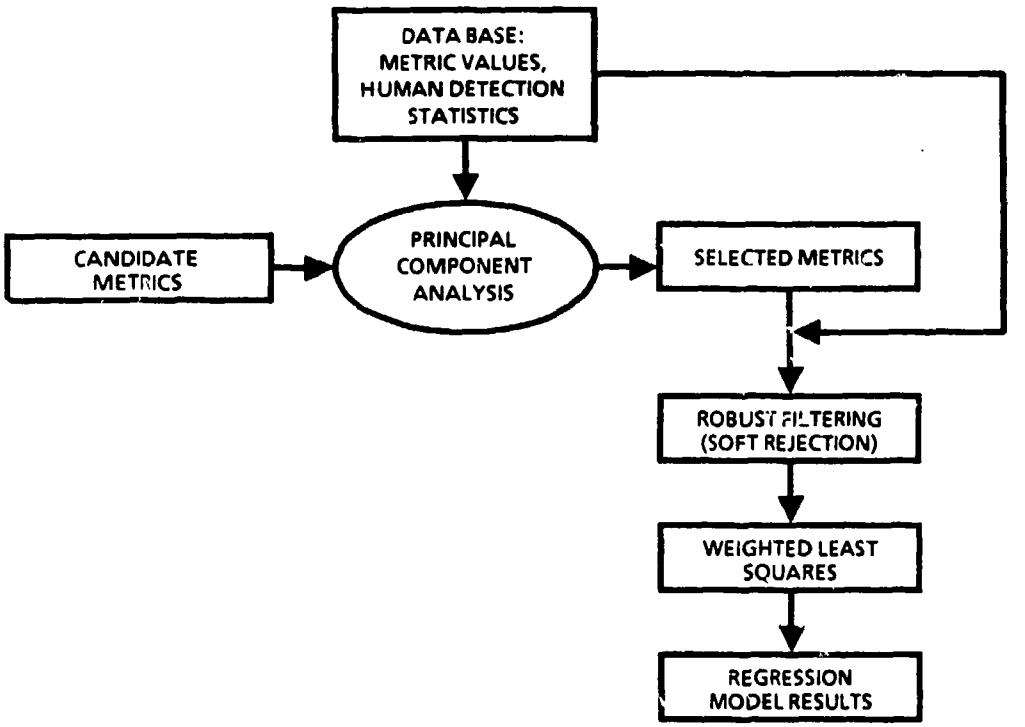


Figure 7. PROCEDURE FOR DETERMINING REGRESSION MODEL FOR PREDICTION OF DETECTION PERFORMANCE AS A FUNCTION OF IMAGE QUALITY

(Because the metrics are correlated to some extent, the coefficients shown do not necessarily indicate the degree or sense of their correlation with P_D.)

Each metric is conditioned on target and background samples that lie in the innermost and outermost portions of a mask that is centered on the target. Two forms of masks were used. One form is a triple-gate mask that was used for grey shades. The other is a double-gate mask that was used for edge statistics. These are shown in Figure 8. The rationale for using two types of masks is as follows. In the case of grey shades, the triple-gate mask was adopted from the earlier work done at NRTC. In the case of edge statistics, it was found that the major edge responses occurred outside the target region and in the guard region. Therefore, the target region was expanded, and the guard region was eliminated.

(a) For grey shades metrics m_1, m_2, and m_4

(b) For edge-based metric m_2

Figure 8. TARGET AND BACKGROUND SAMPLING

5. IMAGE METRIC VALIDATION

The objectives of the image metric validation are composed of three parts: (1) validate TIR and CTRE approaches for prediction/evaluation of the Haar region of interest operator; (2) detect any flaws or weaknesses in TIR and CTRE; and (3) modify/select the most promising approach.

The primary intent of the validation effort is to determine whether anything is basically wrong with TIR (or any of its alternate forms) or CTRE, in terms of correlating with the Haar region of interest operator. Precise values of functions are not needed for this purpose. Therefore, in the interest of efficiency, approximations were made concerning the matching of image metric mask sizes with target sizes, sampling of the image data base, and the form of the correlation metric. The approximations were adequate to determine basic trends and the relative merits of image metrics. (When Project Task 2.2, Characterize Data, is performed, the full treatment for variable target sizes and shapes will be carried out, and more extensive sampling of the image data base will be conducted.)

The underlying assumption in correlating results from different algorithms (including image metrics and region of interest operators) is that the information being extracted by the algorithms is correlated. In the case of military ground vehicle targets, this may not be an unreasonable assumption. However, in the case of natural clutter, there is much more variety of types of objects, as well as variability within each type, compared with military ground vehicle targets. One would thus expect a lower degree of correlation, if any, between algorithms for natural clutter than for targets.

In this major section of the report we first present our approach in Section 5.1. Here we describe our basic approach (Section 5.1.1) which consists of viewing the problem very generally. This includes the rendering of Haar outputs and image metrics to a common reference before attempting a correlation analysis. Two types of correlation analyses are defined (a priori and a posteriori); and two types of data (raw and smoothed) are defined.

Next in Section 5.1.2 is described the imagery selected for the validation exercise. It consists of two data subsets. One for large targets; the other for small targets. There

are twenty-two images for the first subset and twenty-eight images for the second subset.

The two types of Haar detection reports that were used for validation of the image metrics are described in Section 5.1.3; they are: (1) the ATC Haar FOM; and (2) the Haar FOM.

The four types of image metrics that were correlated against the two types of Haar detection reports are described in Section 5.1.4. These consist of: (1) TIR, (2) TIR_{new}, (3) CONTRAST, and (4) CTRE.

The method that was used to register Haar detection reports with image metric reports is described in Section 5.1.5. And the technique for mathematically correlating Haar reports with image metrics is described in Section 5.1.6.

The results of correlating the two types of Haar detections with the four types of image metrics, for both raw and smoothed measurements, for both a priori and a posteriori conditions, and for both large and small targets are presented in Section 5.2.

5.1 Approach

The approach we have taken is described in six parts: (1) Basic Approach; (2) Imagery Selection; (3) Haar Detection Reports; (4) Image Metric Reports; (5) Registration of Reports; (6) and Correlation Analysis.

5.1.1 Basic Approach

Our basic approach is to determine to what extent various candidate imagery metrics correlate with the Haar region of interest operator. This requires a rendering of the responses of the region of interest operator and the imagery metrics to a common reference, both spatially and in magnitude. Once this is accomplished, the two responses may be correlated. The type of correlation required is a positive monotonic relation; in other words, as the magnitude of the Haar response increases, the magnitude of the image metric also increases. The frequency with which this occurs over a number of samples is a measure of the degree of correlation.

Two points of view may be taken of the correlation analysis. One is to consider the sample set as the set of actual detection reports. This can be called the a posteriori correlation–that is, after detection. The other is to consider the sample set as the set of all possible detection reports. This can be called the a priori correlation–that is, prior to detection.

The utility of each point of view is as follows: The a posteriori viewpoint is useful for evaluation of the reliability of detection reports, including ranking of detections by a detector figure of merit (FOM). The a priori viewpoint is useful for evaluation and/or prediction of detection performance, for example, in terms of an operating characteristic curve (See Figure 9). These viewpoints are consistent with the current descriptions of "probability of detection report reliability" and "probability of detection" advanced by the Automatic Target Recognizer Working Group (ATRWG) "Target Recognizer Definitions and Performance Measures" (Reference 5). Both viewpoints are treated in the correlation analysis reported herein.

The most common use of image signal quality measures will be to develop operating curves showing ATR detection performance for signal quality levels above some minimum. Such information is useful for sensor system designers and system analysts. This implies that integral measures such as probability of detection and probability of false alarm for image signal quality levels above some specified threshold value are of most interest, rather than raw density measures such as detectability of individual objects for a specific image signal quality value. In this work we refer to the integral measures as smoothed values and the raw density measures as raw values.

5.1.2 Imagery Selection

To achieve the objectives described above, it was unnecessary to process and analyze our entire data base. To minimize required processing time, it was decided that a single size filter would be used for each frame, although it would be scaled from one frame to another frame. A survey of the available imagery, Texas Instruments (TI) imagery, supplied by the Environmental Research Institute of Michigan (ERIM), indicated that there were many images such that, within a given image, the

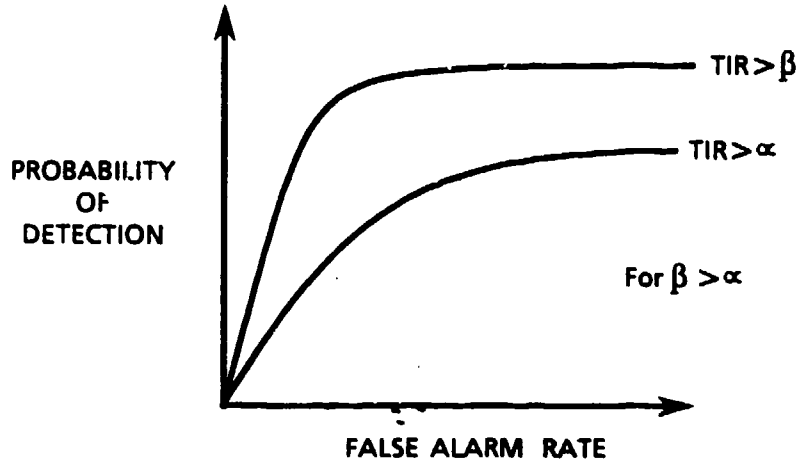


Figure 9. TYPICAL CHARACTERISTIC CURVES

maximum difference in target vertical extent was less than a few rows and the maximum difference in horizontal extent was less than few columns.

We first partitioned the data into a subset (later referred to as the Large Target Validation Data Base) in which for each image selected, the following criteria were met:

(1) Targets within each image do not vary in size by more than five pixel rows in height and five pixel columns in width.

(2) At least one target in each image has an area between 100 pixels and 600 pixels.

The area for each target is obtained from word T-34 of the ERIM target metrics appended to the ATRWG header of each image (Reference 6). Word T-31 of the ERIM target metrics was used to determine target height (vertical), and word T-32 was used to determine target width (horizontal).

The resulting Large Target Validation Data Base consists of the following 22 images (Table 1):

Table 1. LARGE TARGET VALIDATION DATA BASE

ERIM Tape Number	ERIM Image ID	Northrop Image ID
3002-11	2009012	202104
3002-11	2018976	202126
3002-11	2010007	205128
3002-11	2003042	203134
3002-11	2008023	205136
3002-11	2027029	202142
3002-11	2007042	203143
3003-11	2010061	201203
3003-11	2009010	203205
3003-11	2010012	202226
3003-11	2012074	203228
3003-11	2008010	203235
3003-11	2007076	202243
3004-11	2027057	203301
3004-11	2007013	202333
3004-11	2003044	202334
3004-11	2014027	203335
3004-11	2017010	203340
3004-11	2007045	201342
3004-11	2007028	202343
3004-11	2027007	205344

This data base contains 63 targets.

To prevent drawing conclusions on the basis of only large targets, a second validation data base was later partitioned for small targets. Each image in this data base was selected according to the following criteria:

(1) Targets within each image do not vary in size by more than three pixel rows in height and five pixel columns in width.

(2) At least one target in each image has an area between 15 pixels and 30 pixels.

The resulting Small Target Validation Data Base consists of the following 28 images (Table 2):

Table 2. SMALL TARGET VALIDATION DATA BASE

ERIM Tape Number	ERIM Image ID	Northrop Image ID
3002-11	2012007	204111
3002-11	2016003	207112
3002-11	2015032	210118
3002-11	2015046	209122
3002-11	2014040	210130
3002-11	2008051	207137
3002-11	2003048	209138
3002-11	2007048	210145
3003-11	2013013	206206
3003-11	2006032	210207
3003-11	2016000	209212
3003-11	2009071	205216
3003-11	2016032	210222
3003-11	2012051	207230
3003-11	2008035	207238
3003-11	2008000	210239
3003-11	2008067	207241
3004-11	2009000	210307

Table 2. SMALL TARGET VALIDATION DATA BASE (Cont'd)

ERIM Tape Number	ERIM Image ID	Northrop Image ID
3004-11	2012032	210312
3004-11	2009032	210317
3004-11	2010019	206329
3004-11	2012048	209330
3004-11	2006003	206337
3004-11	2008048	209330
3004-11	2008016	210339
3004-11	2017000	210341
3004-11	2007019	207345

This data base contains 83 targets.

5.1.3 Haar Detection Reports

As mentioned in Section 5.1.1, correlation of image metrics with the Haar region of interest operator requires a rendering of the Haar responses to a common reference with the image metrics. All the image metrics (except for CONTRAST) are invariant with sensor system gain and offset. They are also local metrics that are independent of large-scale scene content. The image metrics are also post-processed to map them from the pixel-level to the object-level (as described in Section 3.3).

Considering the correlation of Haar responses with image metrics, we first of all note that the ATC Haar output is not a measure. It is an event: detection or no detection. However, prior to making the detection decision a continuous-valued measure or figure of merit exists; this is the raw Haar filter output. (See Appendix E.) Unfortunately, as the sensor system gain changes, a corresponding change results in the raw filter output. To produce an Haar output that is gain and offset invariant, like the image metrics, a normalization is performed yielding the Haar Figure of Merit (HFOM). (The HFOM is described in Appendix F.) The HFOM is defined at the pixel level. So some post-processing is required to produce reports at the object level.

In these analyses two types of Haar outputs were used to correlate against image metrics. One type was the current ATC Haar output, with a figure of merit appended, referred to as ATC Haar FOM. In this case, the detection decision is based only on the current ATC Haar logic. ATC Haar thresholding logic is based on a percentile threshold that varies with estimated range. (See Appendix E for details.) It determines the region for which an HFOM value is to be obtained. The other type of Haar output to be used for correlation with image metrics was the result of using the HFOM value as the detection variable; this type of output is referred to as Haar FOM. In this case the HFOM is computed at every pixel to form an HFOM map. A simple threshold set at HFOM = -0.30 is then applied to determine pixels of interest. In either case, these are then post-processed by applying a 3 x 3 max operator (corresponding to 6x6 in the original image) to obtain the local maximum HFOM value to be used in the subsequent correlation analysis.

For both types of Haar output, moderately low detection thresholds were used. Thresholding was performed for two reasons: (1) to limit the number of samples so that the number of sample detections would not be computationally excessive; and (2) to restrict the problem to an interesting domain in which most targets are detected, without being swamped with false alarms. No significant effect on the correlation analyses should result from this thresholding. Over the Large Target Data Base and the Small Target Data Base, the thresholding resulted in probabilities of detection of 92% and 70%, respectively, for ATC Haar, and 98% and 78%, respectively, for Haar FOM. Table 3 shows examples of Haar detection lists for the same frame.

5.1.4 Image Metric Reports

Initially four image metrics were investigated: TIR, TIR_{new}, CONTRAST, and CTRE (P_D). These were defined in Eqs. (2-6), (2-8), (2-9), and (4-1).

For any one of the non-CTRE metrics, the way in which the metric value was obtained for this validation exercise is the same as that described in Section 3.2 and 3.3, except that only a single mask was used for each image, although it was scaled in size from one image to the next (as described later). For CTRE, the mask used for

TABLE 3

Detection Reports for ERIM Frame No. 2007028
(a posteriori Analysis)

ATC Haar Detections with Attached FOM					Simple Threshold Detections on HFOM				
Centroid (x_c,y_c)	Expected Width	Height	FOM	True/False Target	Centroid (x_c,y_c)	Expected Width	Height	FOM	True/False Target
(164,94)	22	22	.33	True	(164,94)	22	22	.33	True
(144,110)	22	22	.28	False	(192,118)	22	22	.35	True
(190,120)	22	22	.35	True	(170,138)	22	22	.34	False
(216,140)	22	22	.41	True	(216,140)	22	22	.41	True
(290,262)	22	22	.27	False	(262,148)	22	22	.31	False
(308,278)	22	22	.29	False	(192,160)	22	22	.34	False

each of the elemental metrics was as described in Section 4. They too were kept constant for a given image, but were scaled in size from one image to the next. For a double-gate mask, the procedure followed in determining the inner gate (or target region) dimensions is as follows. For a given image, the target with the maximum width (known from ground truth) was chosen as the reference target. The inner gate dimensions were then scaled to match those of the circumscribing rectangle about the reference target. For the outer gate (or background region), the dimensions were made equal to the integer value closest to $\sqrt{2}$ times each dimension of the inner gate. Thus, the number of pixels in the outer gate ($\sqrt{2}\,\ell \cdot \sqrt{2}\,w - \ell w = \ell w$) is approximately equal to the number of pixels in the inner gate.

By choosing the inner gate to correspond to larger (often, the largest) target in a given image, we intended to minimize the error in estimating TIR, while utilizing a single mask per image. (See Appendix C concerning the larger TIR error associated with contaminating the background region with target pixels.)

For CTRE metrics that use a triple-gate mask, the outer dimensions of the guard region were chosen to match the reference target's circumscribing rectangle; that is, $H_G = h$ and $L_G = \ell$. The innermost (or target) region dimensions were then chosen to be $H_I = H_G\text{-}2$ and $L_I = L_G\text{-}2$. The outermost (or background) region dimensions were chosen to yield an area equal to the target region, or $H_oL_o - H_GL_G = H_IL_I$, such that H_o exceeds H_G and L_o exceeds L_G by a fixed amount, x; x can be determined by substitution and application of the quadratic formula.

An analysis was made of the relative differences in the areas of circumscribing rectangles about targets with respect to the area of the reference rectangle. The results are shown below in Table 4.

From this analysis one can see that the typical relative error in area is less than – 23%. From the sensitivity analysis in Appendix C, one would expect that the typical error in TIR would be biased downward by less than 23%. This was deemed acceptable. However, in the case of the two small targets that were larger than their references, their TIR values may be considerably more depressed below their true values. This could produce significant errors in the correlation analyses, therefore, for purposes of this validation exercise, the images containing these two targets were

eliminated from the correlation analysis. These were ERIM Image I.D. Numbers 2016003 and 2016032.

TABLE 4. RELATIVE DIFFERENCES IN TARGET AREAS WITH RESPECT TO REFERENCE TARGETS

Data Base / Size of Rectangle	Large Targets		Small Targets	
	Number	Average Relative Difference in Area	Number	Average Relative Difference in Area
Smaller than Reference	59	- 0.16	81	- 0.23
Larger than Reference	4	+ 0.15	2	+ 0.50

For each metric, after generating the pixel-level metric values, post-processing is performed as previously described in Section 3.3. From this, a list of object-level metric values is produced consisting of all plateaus that contain their own maximum source point. As in the case of the Haar outputs, a low image metric threshold is imposed on the raw metric value to limit the number of reports while still including weak targets. (In fact, 100% of the targets were included.) This was done for the a priori analyses. In the case of the Large Target Validation set, the thresholds for TIR and CTRE were set at 1.07 and 1.03, respectively; and for the Small Target Validation set, they were set at 0.82 and 1.01, respectively.

Table 5 shows examples of TIR and CTRE lists for the same frame.

5.1.5 Registration of Haar Detection Reports with Image Metric Reports

In the case of targets, to associate Haar detection reports with image metric reports, use was made of the ground truth location of each target. Given its location (x_t,y_t) a search region of $\{(x,y)\}$, such that

TABLE 5

Image Metric Reports for ERIM Frame No. 2007028
(a priori analysis)

TIR Report					CTRE Report				
Centroid (x,y)	Ground Truth Width	Ground Truth Height	TIR Value	True/False Target	Centroid (x,y)	Ground Truth Width	Ground Truth Height	CTRE Value	True/False Target
(371,21)	21	15	1.28	False	(98,42)	21	15	1.04	False
(186,55)	21	15	1.21	False	(186,55)	21	15	1.09	False
(166,98)	21	15	1.86	True	(165,66)	21	15	1.03	False
(186,124)	21	15	1.64	True	(165,96)	21	15	1.28	False
(212,144)	21	15	1.07	True	(186,124)	21	15	1.26	True
(182,304)	21	15	1.11	False	(195,137)	21	15	1.05	True
					(211,144)	21	15	1.04	False
					(287,149)	21	15	1.07	False

$$\left(x_t - \frac{w}{2} \le x \le x_t + \frac{w}{2}\right) \tag{5-1}$$

and

$$\left(y_t - \frac{h}{2} \le y \le y_t + \frac{h}{2}\right) \tag{5-2}$$

is defined in which the largest image metric value is used and the largest HFOM is used. (Targets whose image metric value is zero are deleted from the correlation analysis since the apparent signal quality is extremely low and no meaningful contribution would be made to the correlation analysis.)

In the case of non-targets, for the a posteriori analysis, the association between Haar detection reports and image metric reports is made for those objects passing their respective prescribed Haar thresholds (previously described). The association procedure followed was to use the Haar centroid, width, and height to define a search region in which the largest image metric value is selected for association with the detected object Haar FOM. For the a priori analysis those objects passing their respective image metric thresholds (previously described) were associated with Haar detection reports by using the reported image metric centroid and the ground truth width and height to define a search region in which the largest Haar FOM is selected for association with the reported image metric value.

In any of these registration procedures, an undetermined error exists in the associations made between detection reports and image metrics. However, because different operators respond differently to a given object and its immediate surroundings, a great deal of effort would be required to unambiguously associate such reports in every instance. And since the search regions are limited by the target size (directly or indirectly) it is not likely that the associations made are between different objects.

5.1.6 Correlation Analysis Method

Kendall's test (Reference 7) was used to determine the correlation between image metrics and Haar detection reports. Kendall's test uses rank order to determine correlation. It is a distribution-free test. It was chosen because it does not require

that a functional form (e.g., linear) relating the two variables be assumed. Kendall's test was applied in the following way.

For a given data base of n associated pairs (u,v) where u is an image metric and v is the HFOM value of either type of Haar detection, the set of u's is arranged in order of increasing value. Each associated v is thereby put into a resultant order. If the resultant order of v's is also strictly increasing, then Kendall's coefficient τ of rank correlation would be +1. If the resultant ordering of v's was strictly decreasing, τ would be -1. And if the resultant ordering of the v's was totally random, meaning no correlation, the expected value of τ is 0.

The value of Kendall's τ is determined by first computing two parameters, T_k and I_k, for each v_k, as ordered by the sort on u_k (where $u_k < u_{k+1}$). T_k is defined for each v_k as the number of occurrences of $v_k < v_\ell$, where $\ell > k$. Likewise, I_k is defined as the number of occurrences of $v_k > v_\ell$, where $\ell > k$.

From the sets of T_k and I_k, the summary scores of

$$T = \sum_{all\ k} T_k \quad , \tag{5-3}$$

and

$$I = \sum_{all\ k} I_k \tag{5-4}$$

are computed. From these, Kendall's τ is determined.

$$\tau = \frac{T - I}{\frac{1}{2} \bullet n \bullet (n-1)} \quad . \tag{5-5}$$

Kendall's τ can be related to the more common correlation coefficient ρ. (See Appendix G.)

So for each set of candidate image metric values and each set of corresponding Haar detection FOM reports, a value for Kendall's τ is obtained, as well as other associated statistics for evaluating the confidence in the estimated values of τ. This is done for each condition, a priori and a posteriori. It was also done separately for targets and non-targets, and for both raw and smoothed measurements. (The smoothed measurements consisted of using all samples whose image metric was above a variable threshold value u_T and the average HFOM $\bar{v}$ of those samples.) Figures 10

and 11 show examples of raw and smoothed measurements. Figure 12 provides a simplified description of the entire process using pseudo code.

5.2 Results

We began our validation exercises with the Large Target Data Base, using all four image metrics. A raw image from this data base (ERIM Frame No. 2007028), corresponding ground truth, image signal quality maps, and corresponding HFOM maps are shown in Figures 3 through 6 and 13 through 22. This was followed by validation exercises with the Small Target Data Base.

Based upon these data, results were obtained and summarized in terms of an estimate of Kendall's τ, the 95% confidence interval C about the estimate, the number of samples n, and α, the probability of obtaining the reported value of τ if the image metric and the Haar output were truly uncorrelated (α is rounded off to the nearest one-hundredth). Appendix H provides a numerical example of how these parameters are determined. The results are presented in Tables 6 through 15.

Except for the a priori correlations of TIR and CTRE with non-targets for the Small Target Validation Data Base shown in Tables 14 and 15, all the available data were utilized to compute the statistics shown in these tables. For the two exceptions cited, the total number of non-target samples exceeded 8000. To pare down this number to a reasonable yet representative sample set, 945 individual samples were selected randomly and used for the correlation analysis.

With a view toward eliminating image metrics that do not correlate well with either type of Haar output, TIR_{NEW} is an apparent candidate for elimination. In Tables 8 and 9, TIR_{NEW} shows a strong tendency to be uncorrelated with both Haar outputs (large values of α and/or small values of τ). Therefore TIR_{NEW} was eliminated immediately after the initial experiments with large targets. Upon consideration of FLIR image phenomena, this is not unexpected. For example, a high contrast target should be easily detected; however, TIR_{NEW} may provide a low value if the target region contains both target and background pixels since σ_T would be large.

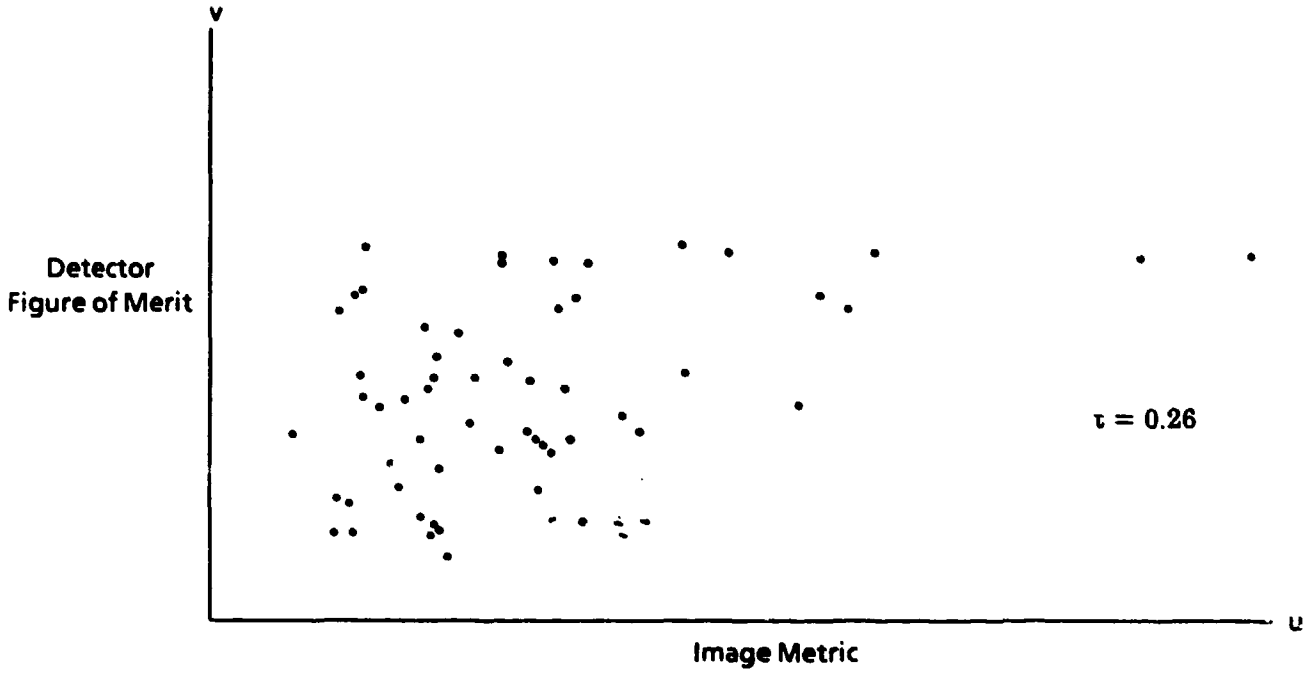


Figure 10. Example of Raw Measurements

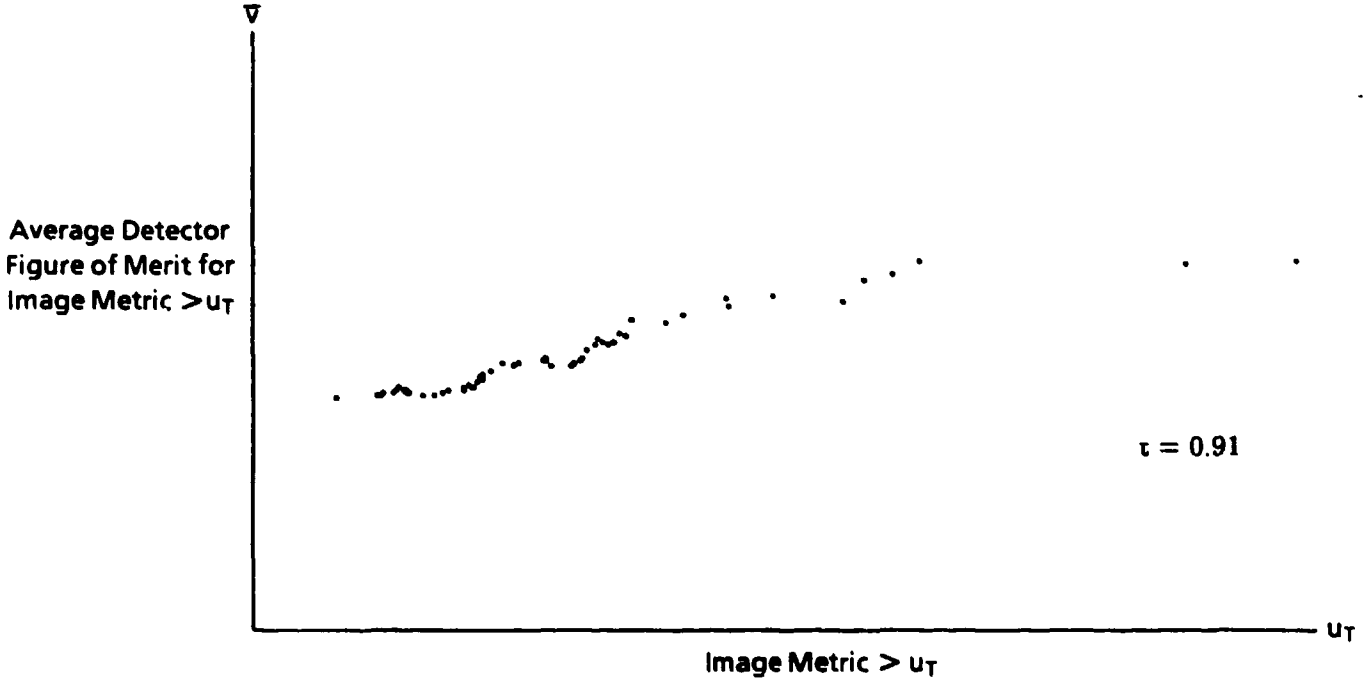
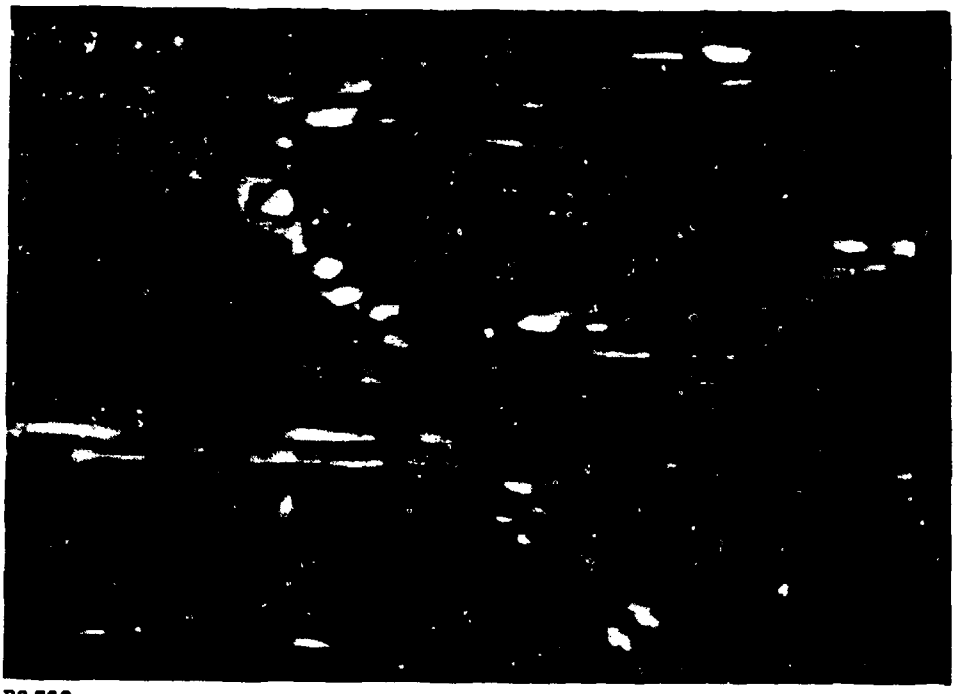


Figure 11. Example of Smoothed Measurements Obtained from Figure 10

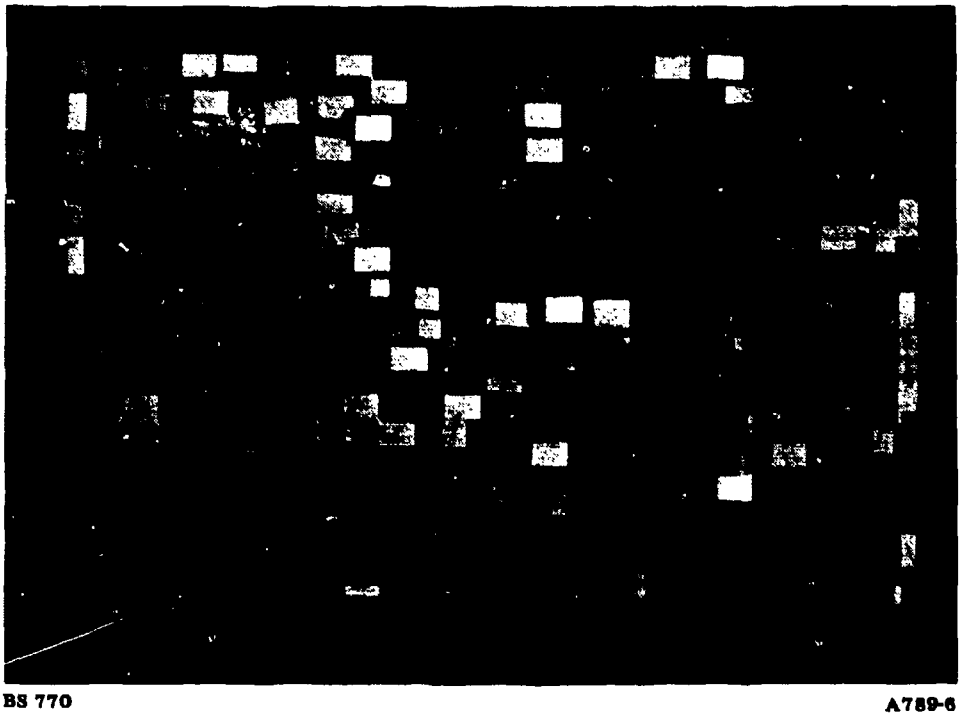
```
START
SELECT MODE (A PRIORI/A POSTERIORI)
SELECT IMAGE
SELECT IMAGE METRIC
SELECT HAAR DETECTION LOGIC
IF MODE IS A PRIORI
   GENERATE LIST OF OBJECTS WITH
      LOW IMAGE METRIC THRESHOLD
   COMPUTE HFOM FOR EACH OBJECT
   COMPUTE KENDALL'S TAU FOR
      RAW VALUES OF METRIC AND HFOM
   COMPUTE KENDALL'S TAU FOR
      SMOOTHED VALUES OF METRIC AND HFOM
ELSE IF MODE IS A POSTERIORI
   GENERATE LIST OF HAAR DETECTIONS
      FOR LOW THRESHOLD
   COMPUTE IMAGE METRIC FOR
      EACH OBJECT
   COMPUTE KENDALL'S TAU FOR
       AW VALUES
   COMPUTE KENDALL'S TAU FOR
      SMOOTHED VALUES
END IF
STOP
```

Figure 12. SIMPLIFIED PSEUDO CODE DESCRIBING PROCESS



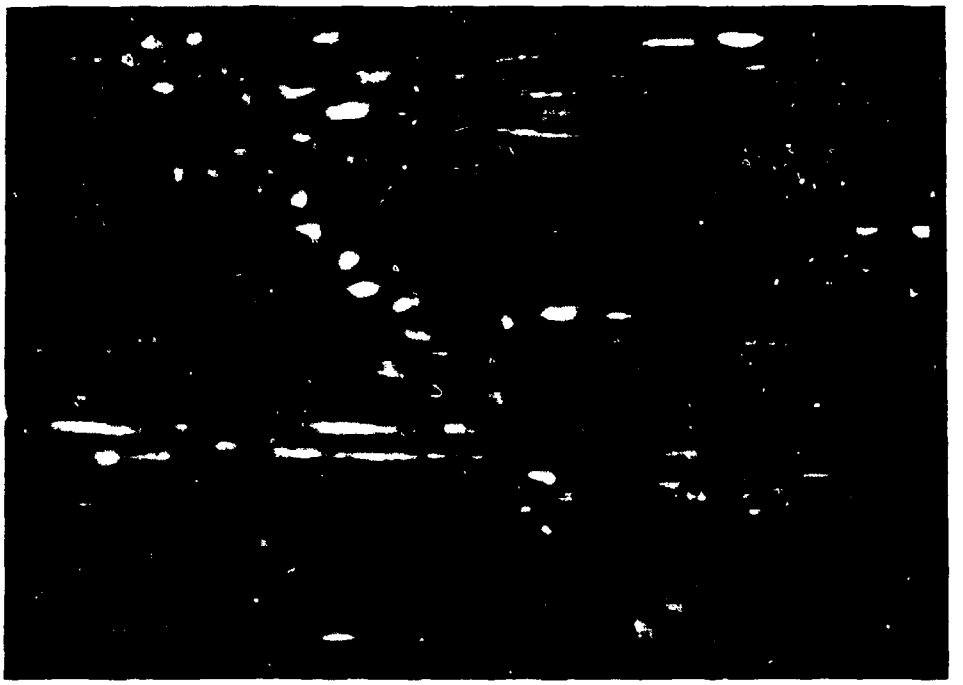
BS 768 A789-5

Figure 13. TIR NEW MAP (EXPONENTIAL SCALING)



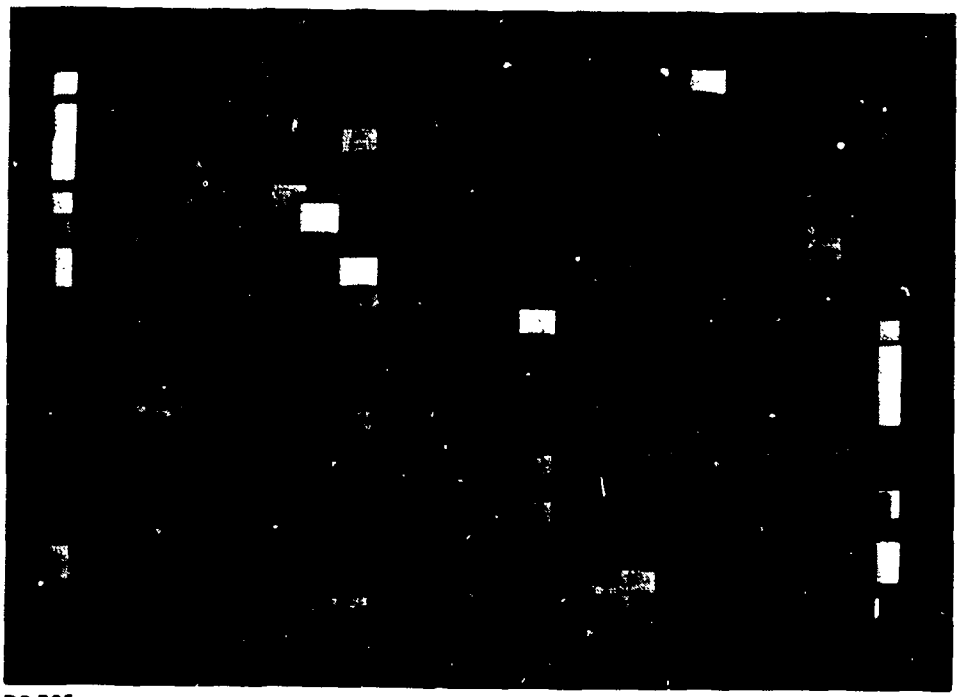
BS 770 A789-6

Figure 14. POST-PROCESSED TIR NEW MAP



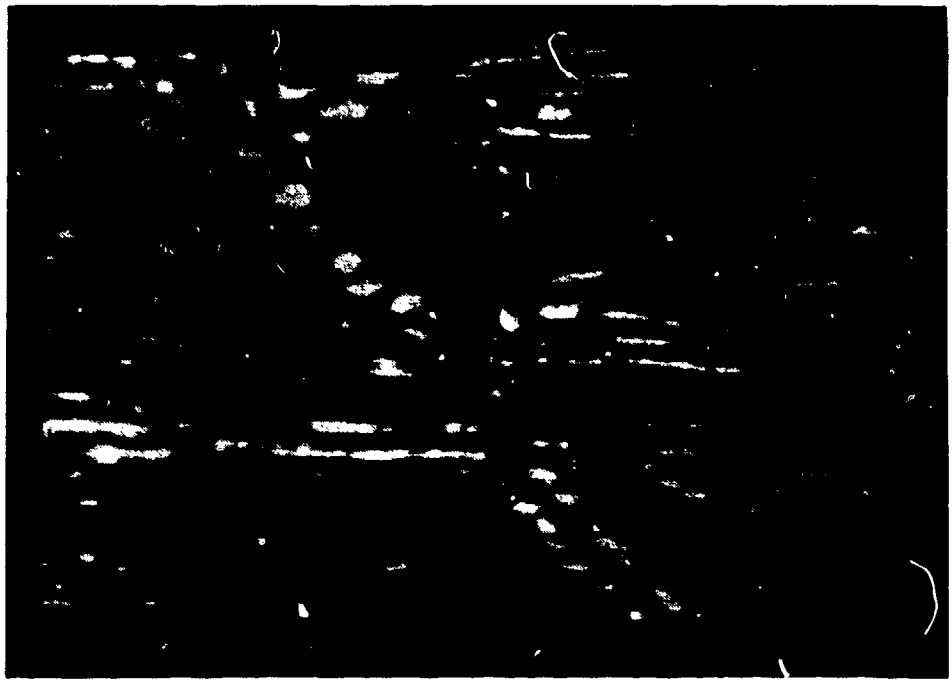
BS 767 A789-7

Figure 15. CONTRAST MAP



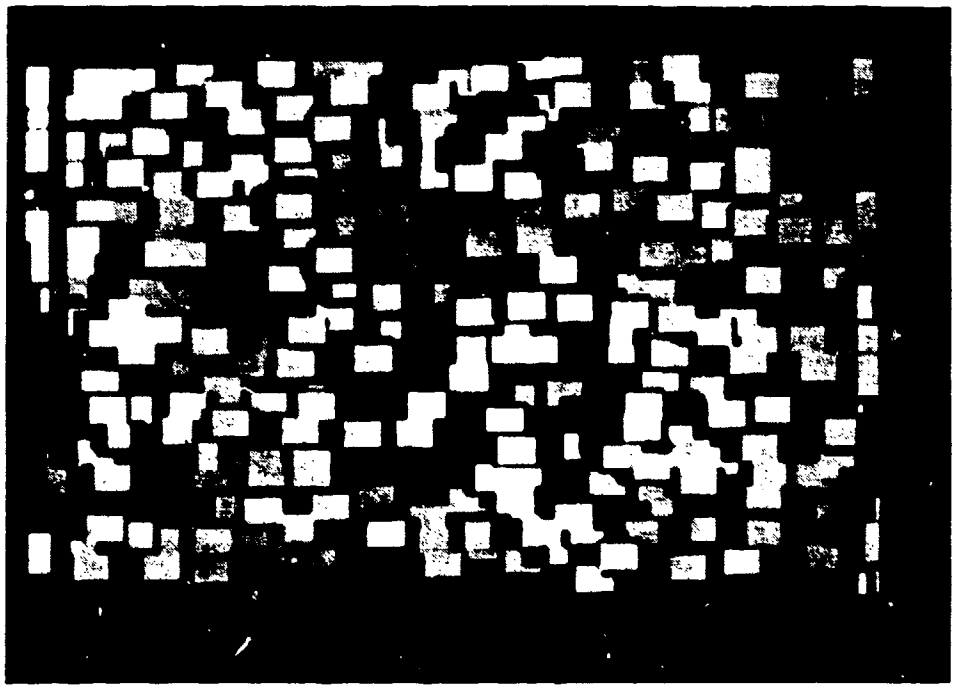
BS 765 A789-8

Figure 16. POST-PROCESSED CONTRAST MAP



BS 774 A789-9

Figure 17. CTRE MAP



BS 772 A789-10

Figure 18. POST-PROCESSED CTRE MAP

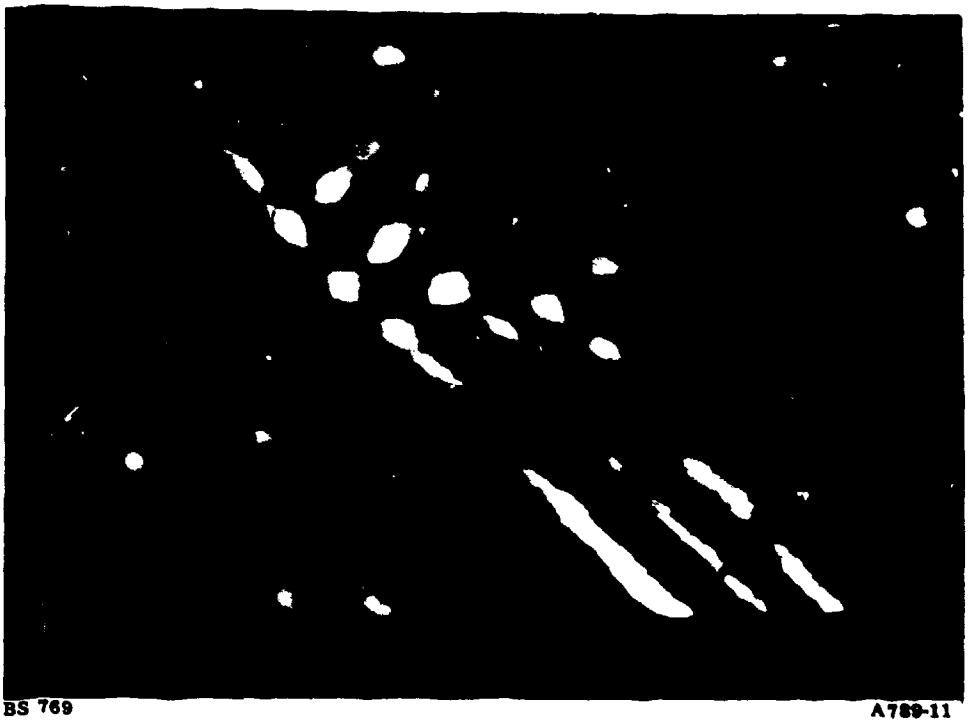


Figure 19. UNNORMALIZED HAAR FILTER MAP

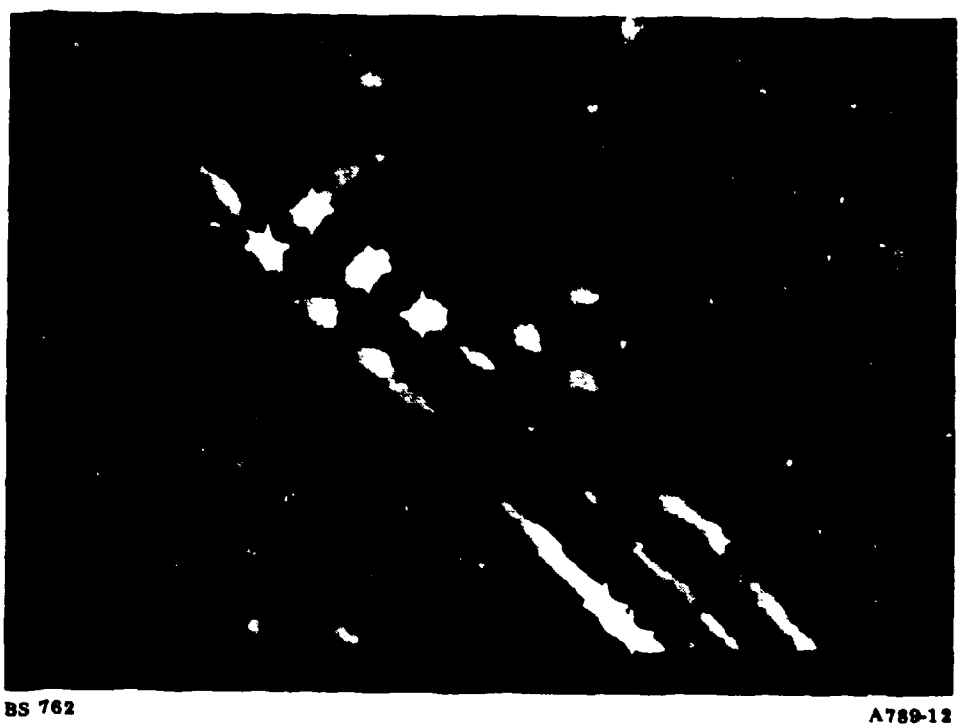


Figure 20. DETECTIONS USING LOW PERCENTILE THRESHOLD ON UNNORMALIZED HAAR MAP

BS 763 A789-13

Figure 21. HAAR FIGURE OF MERIT (HFOM) MAP

BS 766 A789-14

Figure 22. DETECTIONS USING LOW THRESHOLD ON HFOM MAP

Table 6. A POSTERIORI CORRELATION OF CONTRAST WITH ATC HAAR FOM

Image Data Base	Object Type	Measure Type	r	c	n	α
Large targets	Non-Targets	Raw	.21	(-.17,.54)	51	.01
Large Targets	Non-Target	Smooth	.53	(.14,.77)	51	0
Large Targets	Target	Raw	.29	(-.08,.58)	57	0
Large Targets	Target	Smooth	.92	(.63,.98)	57	0

Table 7. A POSTERIORI CORRELATION OF CONTRAST WITH HAAR FOM

Image Data Base	Object Type	Measure Type	r	c	n	α
Large targets	Non-Targets	Raw	.03	(-.37,.42)	42	.40
Large Targets	Non-Target	Smooth	.41	(-.02,.71)	42	0
Large Targets	Target	Raw	.31	(-.05,.60)	61	0
Large Targets	Target	Smooth	.92	(-.65,.98)	61	0

Table 8. A POSTERIORI CORRELATION OF TIR_{NEW} WITH ATC HAAR FOM

Image Data Base	Object Type	Measure Type	τ	c	n	α
Large targets	Non-Targets	Raw	.09	(-.28,.43)	53	.17
Large Targets	Non-Target	Smooth	.26	(-.12,.58)	53	0
Large Targets	Target	Raw	-.23	(-.54,.13)	57	.99
Large Targets	Target	Smooth	-.83	(-.95,.51)	57	1.00

Table 9. A POSTERIORI CORRELATION OF TIR_{NEW} WITH HAAR FOM

Image Data Base	Object Type	Measure Type	τ	c	n	α
Large targets	Non-Targets	Raw	-.075	(-.46,.33)	42	.76
Large Targets	Non-Target	Smooth	-.05	(-.35,.44)	42	.66
Large Targets	Target	Raw	-.18	(-.49,.17)	61	.98
Large Targets	Target	Smooth	-.82	(-.94,-.52)	61	1.00

Table 10. A POSTERIORI CORRELATION OF TIR WITH ATC HAAR FOM

Image Data Base	Object Type	Measure Type	τ	c	n	α
Large targets	Non-Targets	Raw	.19	(-.19,.51)	53	.03
Large Targets	Non-Target	Smooth	.43	(.05,.70)	53	0
Large Targets	Target	Raw	.26	(-.10,.57)	57	0
Large Targets	Target	Smooth	.91	(.42,.98)	57	0
Small Targets	Non-Target	Raw	.12	(-.02,.24)	451	0
Small Targets	Non-Target	Smooth	.92	(.88,.97)	451	0
Small Targets	Target	Raw	.46	(.08,.73)	52	0
Small Targets	Target	Smooth	.97	(.70,1.0)	52	0

Table 11. A POSTERIORI CORRELATION OF TIR WITH HAAR FOM

Image Data Base	Object Type	Measure Type	τ	c	n	α
Large targets	Non-Targets	Raw	.19	(-.23,.55)	42	0
Large Targets	Non-Target	Smooth	.79	(.39,.94)	42	0
Large Targets	Target	Raw	.28	(-.07,.57)	61	0
Large Targets	Target	Smooth	.92	(.64,.98)	61	0
Small Targets	Non-Target	Raw	.13	(-.02,.27)	325	0
Small Targets	Non-Target	Smooth	.86	(.76,.92)	325	0
Small Targets	Target	Raw	.50	(.12,.75)	53	0
Small Targets	Target	Smooth	.98	(.71,1.0)	53	0

Table 12. A POSTERIORI CORRELATION OF CTRE WITH ATC HAAR FOM

Image Data Base	Object Type	Measure Type	τ	c	n	α
Large targets	Non-Targets	Raw	.28	(-.11,.60)	49	0
Large Targets	Non-Target	Smooth	.30	(-.09,.62)	49	0
Large Targets	Target	Raw	.27	(-.09,.57)	57	0
Large Targets	Target	Smooth	.92	(.64,.98)	57	0
Small Targets	Non-Target	Raw	.06	(-.06,.19)	452	.02
Small Targets	Non-Target	Smooth	.91	(.84,.95)	452	0
Small Targets	Target	Raw	.42	(.03,.70)	52	0
Small Targets	Target	Smooth	.92	(.62,.99)	52	0

Table 13. A POSTERIORI CORRELATION OF CTRE WITH HAAR FOM

Image Data Base	Object Type	Measure Type	τ	c	n	α
Large targets	Non-Targets	Raw	.32	(-.12,.65)	40	0
Large Targets	Non-Target	Smooth	.89	(.52,.98)	40	0
Large Targets	Target	Raw	.29	(-.07,.58)	61	0
Large Targets	Target	Smooth	.92	(.65,.98)	61	0
Small Targets	Non-Target	Raw	.14	(.02,.28)	324	0
Small Targets	Non-Target	Smooth	.87	(.77,.93)	324	0
Small Targets	Target	Raw	.46	(.08,.72)	53	0
Small Targets	Target	Smooth	.92	(.62,.98)	53	0

Table 14. A PRIORI CORRELATION OF TIR WITH HAAR FOM

Image Data Base	Object Type	Measure Type	τ	c	n	α
Large targets	Non-Targets	Raw	.19	(.05,.32)	409	0
Large Targets	Non-Target	Smooth	.98	(.94,1.0)	409	0
Large Targets	Target	Paw	.31	(-.05,.59)	62	0
Large Targets	Target	Smooth	.92	(.65,.98)	62	0
Small Targets	Non-Target	Raw	.02	(-.06,.11)	945	.15
Small Targets	Non-Target	Smooth	.78	(.72,.83)	945	0
Small Targets	Target	Raw	.33	(.08,.60)	73	0
Small Targets	Target	Smooth	.96	(.75,1.0)	73	0

Table 15. A PRIORI CORRELATION OF CTRE WITH HAAR FOM

Image Data Base	Object Type	Measure Type	τ	c	n	α
Large targets	Non-Targets	Raw	.11	(.01,.21)	775	0
Large Targets	Non-Target	Smooth	.97	(.93,.99)	775	0
Large Targets	Target	Raw	.29	(-.06,.58)	62	0
Large Targets	Target	Smooth	.91	(.65,.98)	62	0
Small Targets	Non-Target	Raw	-.02	(-.11,.06)	945	.84
Small Targets	Non-Target	Smooth	.35	(.26,.43)	945	0
Small Targets	Target	Raw	.35	(.03,.61)	74	0
Small Targets	Target	Smooth	.93	(.70,.99)	74	0

Also, reviewing the results for the CONTRAST metric for large targets, a posteriori condition, it appears that the values of τ in Table 7 tend to be lower for non-targets than are the corresponding values for TIR and CTRE in Tables 11 and 13. Also the value of α of 0.40 in Table 7 indicates some degree of uncorrelatedness with the Haar FOM for non-targets. In addition, consideration of the fact that CONTRAST is not independent of the FLIR gain determined that it too should be eliminated.

After eliminating TIR_{NEW} and CONTRAST, only TIR and CTRE were kept to carry out the remaining analyses. No significant differences in τ or α are apparent in all of the experiments performed. In fact, for both image metrics, all the smoothed target results for τ lie in the range of 0.91 to 0.98, an excellent correlation. Smoothed non-target results for τ lie in the range of 0.30 to 0.92. (If the anomalously low values of τ obtained for both TIR and CTRE against the same Haar outputs in Tables 10 and 12, namely $\tau = 0.43$ and $\tau = 0.30$, were eliminated the range would become 0.79 to 0.92.) Although the non-target correlations are not nearly as high as for targets, this result is not totally unexpected since the Haar detector and the image metrics were designed with targets in mind, whereas clutter (i.e., non-targets) can be quite varied and therefore can produce quite varied responses to different operators that tend to emphasize different types of image structure.

A strong positive correlation has been demonstrated for both TIR and CTRE, on the one hand, versus both ATC Haar FOM and Haar FOM, on the other hand. Thus, both TIR and CTRE have been validated.

6. CONCLUSIONS AND RECOMMENDATIONS

A number of approximations and constraints have been imposed on the data and there is some degree of ambiguity in the registration process, as noted in the discussion. However, it is expected that the effect of these imperfections on the results of the validation exercise is not large and would not affect the general conclusions reached regarding the validation of the TIR and CTRE metrics nor their relative merit.

In general TIR correlated somewhat better with Haar than did CTRE. (The alternate forms of TIR did not correlate well.) For correlations with targets only, the average TIR correlation was 0.94; and the average CTRE correlation was 0.92. For correlations against non-targets only, the average TIR correlation was 0.79; and the average CTRE correlation was 0.72. On a single-experiment basis, the differences between TIR and CTRE correlation values were not significant, except for one case. If one had to make a choice between TIR and CTRE, TIR should be accepted, primarily on the basis of lesser computational complexity.

As is evident from these correlation values, a useful relationship has been demonstrated for both TIR and CTRE as both predictors and explainers of Haar detection performance. The aggregate results do not appear to be very sensitive to the precise shape and size of the operator masks used.

Based on these experiments, as well as prior work in this area, we expect that the major objective of this effort has been achieved; namely, a means has been provided and validated for quantitative measurement of signal quality (for both targets and clutter) for use in ATR evaluation.

Recommendations for future work on image metrics that builds upon this effort, as well as other related efforts are listed below: (1) To reduce the computational burden of computing the TIR values over several masks for every pixel in an image, investigations should be performed to determine the errors introduced through various approaches for image subsampling; (2) To better determine the appropriate mask dimensions for realistic sensors, investigations should be performed of the two-dimensional target response to point spread functions, as well as to TIR masks; (3) For purposes of comparative analysis, the convolutional TIR values used in this

work should be compared with the wire-frame TIR used by ERIM; and (4) For evaluation of ATR classifiers new metrics need to be established and validated, perhaps involving the number of resolution cells on target and some edge statistics.

REFERENCES

1. Y. S. Hsu, "Target to Interference Ratio Definitions," NRTC 330-112-80, 1 August 1980.

2. Y. S. Hsu, "TIR Measure Results," NRTC 330-09-81, 20 February 1981.

3. J. J. Reis and J. A. Reed, "Advanced Sensor Signal Processor (U)," Final Technical Report, June 1984 (Confidential). (Distribution limited to U.S. Government agencies only through Commander, U.S. Army Missile Command, Attn: DRSMI-0).

4. G. B. Silverman, "Criteria for Target Recognizer Evaluation," ATLPD-85-Y73-001, 8 November 1985.

5. "Target Recognizer Definitions and Performance Measures," ATRWG 86-001, January 1986.

6. "Image Characterization, Truthing, and Data Exchange Formats for AIRT/CAPIR: Code Word Definitions for Local and Global Metrics," ERIM SMT 85161 (revised 11 Oct. 1985).

7. M. G. Kendall, Rank Correlation Methods, Charles Griffin Co., Ltd., London 1962.

APPENDIX A

Estimate of σ_B for TIR When Limited by Quantization Step Size

This appendix addresses the problem of determining TIR when σ_B is computed to be zero. TIR is defined as:

$$TIR = |\mu_T - \mu_B|/\sigma_B \tag{A-1}$$

where μ_T is the computed mean of target window, μ_B is the computed mean of background window, and σ_B is the computed standard deviation of the background.

The above definition of TIR approaches infinity as σ_B approaches zero. In fact, because the grey level values assume only discrete values, σ_B can equal zero.

The problem is to estimate the true value of σ_B when it is computed as zero. Due to quantization error, this value should be within the range [0,1]. We are interested in deriving the upper bound of σ_B because we want to have a conservative calculation of TIR (a large value of σ_B gives a small value of TIR).

Denote δ as the quantization step size, this quantization step size is related to the number of bits used. Assume that within this quantization step, the background grey level has a Gaussian distribution with mean μ_B and standard deviation σ_B. For a Gaussian distribution, σ_B will be the largest if μ_B is at the center of the quantization step. Assume that 99.73% of the population lies within this quantization step, then:

$$3\sigma_B = 0.5\delta \quad \text{or} \quad \sigma_B = 0.167\delta.$$

Alternatively, for a uniform distribution, we have: $\sigma_B = \delta/12 = 0.083\delta$.

If we assume the variance follows a X^2 distribution, then for a 99% confidence interval, we have:

$$\frac{(n-1)\sigma_B^2}{n-1+2.57\sqrt{2(n-1)}} \le \sigma \le \frac{(n-1)\sigma_B^2}{n-1-2.57\sqrt{2(n-1)}}, \text{ for large } n. \tag{A-2}$$

$$\frac{(n-1)\sigma_B^2}{X^2} \leq \sigma \leq \frac{(n-1)\sigma_B^2}{X^2} \text{ , for small } n. \qquad \text{(A-3)}$$

Typical values of the range of the estimate for σ_B are:

Population n	Confidence Interval (%)	Type of Distribution	Range
30	90	Gaussian	[0.1357, 0.2087]
50	99	Gaussian	[0.1352, 0.2404]
100	99	Gaussian	[0.1426, 0.2092]
30	90	Uniform	[0.0676, 0.1039]
50	99	Uniform	[0.0673, 0.1197]
100	99	Uniform	[0.0710, 0.1042]

From the above, we can take the conservative number (i.e., the largest number of σ_B) for σ_B as 0.2404 or approximately 0.25 δ.

APPENDIX B

TIR as a Test for Zero Contrast

Below is derived a statistical test of the hypothesis of zero contrast between target and background region (i.e., their means are equal) vs. the alternative hypothesis of a non-zero contrast. What is of interest is that the derived test statistic is of a form very similar to TIR. Thus TIR can be thought of as a test for non-zero contrast. In deriving the test statistic it is assumed that the data are Gaussianly distributed. However, it will be pointed out afterwards that the test is aymptotically distribution-free under very general conditions (which always apply in the case of FLIR):

Let $(x_1,x_2,...,x_n)$ be the pixels from the target and let $(y_1,y_2,...,y_m)$ be the pixels from the background. Assume both sets of pixels are independently and identically drawn from their respective Gaussian populations, with means and variances of (m_t,v_t) and (m_b,v_b), respectively. Then, let $z_i = x_i - y_i$. The average value of z, $\bar{z}$, is $\bar{z} = \bar{x} - \bar{y}$. Since z_i is a l inear combination of uncorrelated data x_i and y_i, the variance of z_i, var(z_i), is the sum of the two variances, var(x_i) and var(y_i). Thus, for z, we have var($\bar{z}$) = var($\bar{x}$)/n + var($\bar{y}$)/m, also, $\bar{z}$ is Gaussianly distributed.

Thus under the null hypothesis of zero contrast ($m_t = m_b$), the probability that $|\bar{x}-\bar{y}|/\sqrt{(v_t/n + v_b/m)}$ (note similarity to TIR) takes on a value greater than, say, t, is the area under the standard normal curve to the left of -t plus the area under the curve to the right of t. Therefore, a large value of $|\bar{x}-\bar{y}|/\sqrt{(v_t/n + v_b/m)}$ is extremely unlikely under the zero contrast hypothesis, and a large value will imply that the contrast $|m_t-m_b|$, is greater than zero. Thus, large values of TIR imply low probability that the contrast between target and background is zero.

Although the above derivation of TIR assumes Gaussian distributions for the two sets of pixels, it can be shown that for any pair of distributions in which both random variables are discrete and have a finite number of possible values, e.g., FLIR pixel data, the quantity $|\bar{x}-\bar{y}|/\sqrt{(v_t/n + v_b/m)}$ has a Student's t distribution, which becomes Gaussianly distributed in the limit, as n and m approach infinity. Furthermore, the Gaussian approximation to the Student's t distribution should be good even for fairly small targets, e.g., greater than 100 pixels.

To calculate the endpoints of a 100*(1-α)% confidence interval for the contrast $|m_t - m_b|$, first estimate the target population with the TIR inner box and the sample estimates X_i and S_i^2 for m_t and var(x); and likewise the outer box and X_o and S_o^2 for m_b and var(y). Then in a table of standard normal deviates look up the value $Z_{\alpha/2}$ which satisfies the inequality, $P(z > z_{\alpha/2}) = \alpha/2$, and substitute it into the following inequality,

$$-Z_{\alpha/2} < \frac{(X_i - X_0) - (m_t - m_b)}{\sqrt{\left(\frac{S_i^2}{n} + \frac{S_0^2}{m}\right)}} < Z_{\alpha/2} \quad , \tag{B-1}$$

Now, solve for the left and right endpoints of the interval containing $(m_t - m_b)$; this results in the following:

$$(X_i - X_0) - Z_{\alpha/2}\sqrt{\left(\frac{S_i^2}{n} + \frac{S_0^2}{m}\right)} < (m_t - m_b) < (X_i - X_0) + Z_{\alpha/2}\sqrt{\left(\frac{S_i^2}{n} + \frac{S_0^2}{m}\right)} \tag{B-2}$$

Since we are more interested in the absolute value of contrast than the signed contrast, we derive from the above results the following inequalities for $|m_t - m_b|$:

$$I_L = (X_i - X_0) - Z_{\alpha/2}\sqrt{\left(\frac{S_i^2}{n} + \frac{S_o^2}{m}\right)}, I_R = (X_i - X_o) + Z_{\alpha/2}\sqrt{\left(\frac{S_i^2}{n} + \frac{S_o^2}{m}\right)} \tag{B-3}$$

$$\min(|I_L|, |I_R|) < |m_t - m_b| < \max(|I_L|, |I_R|).$$

Thus, we can say, with only a 100*α% probability of being wrong, that the true contrast between the inner and outer boxes is in the above interval.

The role of the numbers of pixels in the two samples, n and m, in the above inequality, are worth nothing. They demonstrate how any estimate based on sampled data can only be known within certain bounds which are a function of the true variance(s) of the sample(s) and the number of samples. As the number of samples increases, the width of the confidence interval decreases and the accuracy of the estimate increases.

APPENDIX C

Sensitivity of TIR to Target Size Estimation Errors

Presented below is an analysis of the sensitivity of the Target-to-Interference Ratio (TIR) image characterization metric to errors in defining the size and shape of a target. A general case formula relating the impact of errors of assigning target pixels to the background region and background pixels to the target region on the estimated TIR is derived. A guard ring is not in the definition of TIR used in the derivation; however, it will be clear from the final formula what effect a guard ring would have.

The TIR is shown to respond in a complex way to small errors ($\approx 10\%$) in defining the target and background regions of the filter. The impact of this sensitivity on algorithm evaluation tasks should be considered carefully.

Background

The TIR of a target is defined by

$$TIR = |\mu_T - \mu_B|/\sigma_B = |\Delta\mu|/\sigma_B \qquad \text{(C-1)}$$

where μ_T and μ_B are the means of the "target" pixels and the "background" pixels, respectively,and σ_B is the standard deviation of the background region. The definition of TIR depends on the definition of the procedure for designating which pixels are target pixels and which are background pixels. Some suggestions made in the past are:

(1) Let the target region be those pixels in the target's bounding box. Let the background regions be those pixels in a concentric rectangle about the target.

(2) Let only those pixels that pass through a ray from the sensor to the physical target be in the target region. Let the background pixels be defined as in (1) above.

(3) Let the target region be those pixels in a rectangle defined by the expected size of the target, centered on the target. Let the background pixels be defined as in (1).

The analysis below applies equally well to all three cases.

Analysis

Notation: N_T, number of pixels in the target region,
N_B, number of pixels in background region,
μ_T, mean grey level of target region,
μ_B, mean grey level of background region,
σ_T, standard deviation of target region grey levels,
σ_B, standard deviation of background region grey levels.

One type of error is including target pixels in the background region. Let $100e_{TB}$ be the percentage of the N_B pixels in the background region which actually belong to the target. Also, let $100e_{BT}$ be the percentage of the N_T target region pixels which actually belong to the background, which is the other type of error. Let the prime notation (') indicate the estimate of that quantity based on the way the pixels are labelled (and mislabelled). For instance, μ'_T is the mean of the target region pixels, including any errors in labelling. Then,

$$E(\mu'_T) = \frac{N_T(1-e_{BT})\mu_T + N_T(e_{BT})\mu_B}{N_T} = (1-e_{BT})\mu_T + e_{BT}\mu_B \qquad \text{(C-2)}$$

$$E(\mu'_B) = \frac{N_B(1-e_{TB})\mu_B + N_B(e_{TB})\mu_T}{N_B} = e_{TB}\mu_T + (1-e_{TB})\mu_B \qquad \text{(C-3)}$$

$$E(|\Delta\mu'|) = (1-e_{TB}-e_{BT})|\Delta\mu| \qquad \text{(C-4)}$$

Thus, the estimate of contrast $E(|\Delta\mu'|)$ is proportional to (and always less than) the true contrast ($|\Delta\mu|$).

Now, to compute $E(\sigma'^2_B)$ let Z_i be a pixel in the background region, let Y_i be a pixel in the background region that actually belongs to the target, and let X_i be a true background pixel. Then

$$E(\sigma_B'^2) = \sum_{i=1}^{N_B} \frac{(Z_i - E(\mu_B'))^2}{N_B} = \frac{\sum_{i=1}^{M_B} (X_i - E(\mu_B'))^2 + \sum_{i=1}^{M_T} (Y_i - E(\mu_B'))^2}{N_B} \quad \text{(C-5)}$$

where $M_B = N_B(1\text{-}e_{TB})$, $M_T = N_B(e_{TB})$.

Then,

$$E(\sigma_B'^2) = \frac{\sum_{i=1}^{M_B} (\varepsilon_i + \mu_B - E(\mu_B'))^2 + \sum_{i=1}^{M_T} (\alpha_i + \mu_T - E(\mu_B'))^2}{N_B} \quad \text{(C-6)}$$

$$E(\sigma_B'^2) = \frac{\sum_{i-1}^{M_B} (\varepsilon_i + \mu_B - (\mu_B + \delta_B))^2 + \sum_{i=1}^{M_T} (\mu_T + \alpha_i - (\mu_T + \delta_T))^2}{N_B} \quad \text{(C-7)}$$

$$E(\sigma_B'^2) = \frac{\sum_{i=1}^{M_B} \varepsilon_i^2 - 2\delta_B \sum_{i=1}^{M_B} \varepsilon_i + M_B \delta_B^2 + \sum_{i=1}^{M_T} \alpha_i^2 - 2\delta_T \sum_{i=1}^{M_T} \alpha_i + M_T \delta_T^2}{N_B} \quad \text{(C-8)}$$

where $\delta_B = |E(\mu'_B) - \mu_B|$, $\delta_T = |E(\mu'_B) - \mu_T|$, $\varepsilon_i = Y_i - \mu_B$, $\alpha_i = X_i - \mu_T$.

Since the ε_i's are deviations of true background pixels about their mean μ_B, and since the α_i's are deviation of the true target pixels about their mean μ_T, the second and fifth terms in the numerator are approximately zero and the first and fourth are approximately the standard deviation of the true background pixels and true target pixels, respectively.

Thus,

$$E(\sigma_B'^2) \simeq \frac{M_B\sigma_B^2 + M_B\delta_B^2 + M_T\sigma_T^2 + M_T\delta_T^2}{N_B} \quad \text{(C-9)}$$

$$E(\sigma_B'^2) \simeq \frac{N_B(1-e_{TB})\sigma_B^2 + N_B(1-e_{TB})\delta_B^2 + N_B(e_{TB})\sigma_T^2 + N_B(e_{TB})\delta_T^2}{N_B} \quad \text{(C-10)}$$

$$E(\sigma_B'^2) \simeq (1-e_{TB})\sigma_B^2 + (1-e_{TB})\delta_B^2 + e_{TB}\sigma_T^2 + e_{TB}\delta_T^2 \quad \text{(C-11)}$$

To simplify δ_B, δ_T, we have

$$\delta_B = |\mu_B - E(\mu'_B)| = |\mu_B - (1-e_{TB})\mu_B - e_{TB}\mu_T| = e_{TB}|\Delta\mu|,$$

$$\delta_T = |\mu_T - E(\mu'_B)| = (1-e_{TB})|\Delta\mu|.$$

Now,

$$E(\sigma_B'^2) = (1-e_{TB})\sigma_B^2 + (e_{TB})\sigma_T^2 + (1-e_{TB})(e_{TB})^2(\Delta\mu)^2 + (1-e_{TB})(e_{TB})^2(\Delta\mu)^2 \quad \text{(C-12)}$$

$$E(\sigma_B'^2) = \sigma_B^2 + e_{TB}(\sigma_T^2 - \sigma_B^2) + (e_{TB} - e_{TB}^2)(\Delta\mu)^2 \quad \text{(C-13)}$$

If we substitute $E(|\Delta\mu'|)$ and $E(\sigma'_B{}^2)$ from Equations (C-4) and (C-13), respectively, into the TIR definition, we have, after multiplying numerator and denominator by $(1/\sigma_B^2)$,

$$E((TIR')^2) = \frac{(1-e_{TB}-e_{BT})^2(TIR)^2}{1+e_{TB}(\sigma_T^2/\sigma_B^2 - 1) + (e_{TB} - e_{TB}^2)(TIR)^2} \quad \text{(C-14)}$$

Examples

Equation (C-14) shows how errors in defining which pixels belong to the target and which belong to the background affect the calculated value of TIR. For example, (A) if the estimated target size exceeds the true target size by 10%, so that 10% of the pixels in the target region actually are part of the background, then $e_{BT} = .1$,

$e_{TB}=0$, and $E((TIR')^2) = (1-e_{BT})^2(TIR)^2 = (.9)^2(TIR)^2$. So, if $(TIR)^2 = 10.$, then $E((TIR')^2) = .81(TIR)^2$, and $E((TIR')^2) = 8.1$. Thus, the true $(TIR)^2$ is underestimated by 19%.

As another example, (B) suppose the estimate of the targets size falls short of the true size by 10%, so that some target pixels are used in the background calculations. Then, $e_{TB} = .1$, $e_{BT}=0$, and

$$E((TIR)^2) = \frac{(.81)(TIR)^2}{1+.1\frac{\sigma_T^2}{\sigma_B^2} - .1 + .09(TIR)^2} = \frac{(.81)(TIR)^2}{.9+(.1)\frac{\sigma_T^2}{\sigma_B^2} + .09(TIR)^2} \quad \text{(C-15)}$$

If $\sigma_T = \sigma_B$, then

$$E((TIR')^2) = \frac{(.81)(TIR)^2}{1+.09(TIR)^2} \quad \text{(C-16)}$$

Thus, if $(TIR)^2 = 10$, then $E((TIR')^2) = 4.3$. In this example, $(TIR)^2$ is underestimated by 67%.

Discussion

The relation between $(TIR)^2$ and $E((TIR')^2)$ is fairly complicated, as can be seen in (C-14). However, when $e_{TB} = 0$ (i.e., when no target pixels contaminate the background) then it simplifies to $E((TIR')^2) = (1-e_{BT})^2 (TIR)^2$. On the other hand, when target pixels do contaminate the background ($e_{TB}>0$), the ratio of target pixel variance to background pixel variance become important, and the whole expression (C-14) is complex. In either case, $E((TIR')^2)$ always underestimates $(TIR)^2$.

Reviewing the examples above, we see that in (A) a target with $(TIR)^2 = 10$ has $E((TIR')^2) = 8.1$ if the target size is overestimated by 10%, and in (B) that the same target has a $E((TIR')^2) = 4.3$ if its size is underestimated by 10%. Thus, in the first case $(TIR)^2$ is underestimated by 19%, while in the second case it is underestimated by 67%. Thus, using TIR as defined in (1) to characterize target signatures is asymmetrically sensitive to errors in underestimating target size vs. overestimating target size.

Finally, the use of a guard ring must be considered. If the target size is estimated in order to construct a TIR filter mask, a guard ring can be used to decrease the percent

contamination of the target and background regions (e_{TB}, e_{BT}). However, if one does not know the actual error in size estimation, one does not know how to construct a guard ring that will effectively improve the TIR.

APPENDIX D

Robust TIR

All proposed definitions of TIR are functions in which the means and variances of the inner and outer gate regions are free variables. Thus, TIR is always sensitive to errors in estimating those variables. The idea of making TIR less sensitive to errors of those kinds of errors, i.e., making TIR more robust, was considered. One approach would be to replace the estimate of the mean with the estimate of the median and replace the estimate of variance (mean of squared deviations about the mean) with the median of absolute deviations about the median (M.A.D.). For problems in which the data can be modeled as having an asymmetrical probability distribution the sample median is a less sensitive estimate of location than the sample mean. However, since multimodal targets are an important type of target signature, and since the median is not robust for multimodal data, this robust approach was rejected. A second approach would be to directly estimate contrast using a robust form of least squares estimation, and using the M.A.D. instead of the variance. That approach would work well in that the TIR estimate of a target would not be biased downward by another bright target in its background, and the TIR estimate of a clutter object region would not be biased upward by a few noise spikes in the inner box. However, it would always give low TIRs to the type of target signature characterized by a few hot pixels on the target's engine and cold pixels on the remainder of the target, and a cold, immediate background, because it would, in effect, kick the bright pixels out of the calculation of the target's mean, thus making the estimated contrast close to zero. Thus, this appproach was rejected. In summary, no acceptable method of making TIR more robust was generated.

APPENDIX E

ATC Firmware Version of the HAAR Detector

This appendix presents an overview of the Haar detector as it is implemented in the firmware of the Automatic Target Cuer (ATC). Figure E-1 is a flow diagram of the detector including its preprocessing algorithms. As illustrated, prior to Haar detection the image undergoes a 2:1 reduction along the row and column spaces resulting in an image that is 1/4 the size of the original. The reduced image is then passed through the appropriate range dependent Haar mask, and pixel values are hard limited to zero (positive clip). The resultant image is histogrammed and a global threshold is determined. After threshold application, the image is blurred and colored. Estimates of the centroids and sizes of potential targets are then passed on to the segmentation module.

The Haar Mask. The function of any target detector is to locate and isolate regions of interest within the scene so as to minimize the amount of data processing required for silhouetting and classification. The Haar detector achieves this function by using a finite impulse response (FIR) matched filter. The expected target signature is a contiguous target-sized area with similar grey levels called a "blob" or "spot." The Haar matched filter is relatively large, especially at near ranges where the mask size can be 31 x 31. The resulting convolution process can monopolize system resources and be the limiting factor of the system's processing bandwidth. Therefore several modifications were incorporated to increase the hardware processing speed. First, as previously stated, the image undergoes a 4:1 reduction in area. Second, the Haar filter was designed to be decomposable, allowing the mask to be broken down into sub-masks whose non-zero coefficients are either plus or minus one. The convolution process can now occur without the need of a time consuming multiplication operation. Figure E-2 demonstrates how the ATC firmware generates an equivalent complex convolution through a series of simplified convolutions using the sub-masks.

The current ATC implementation has incorporated range into the Haar detector, utilizing four separate Haar convolutional masks designed to operate at ranges from 1 km to 8 km. A single mask is selected for each image based on its

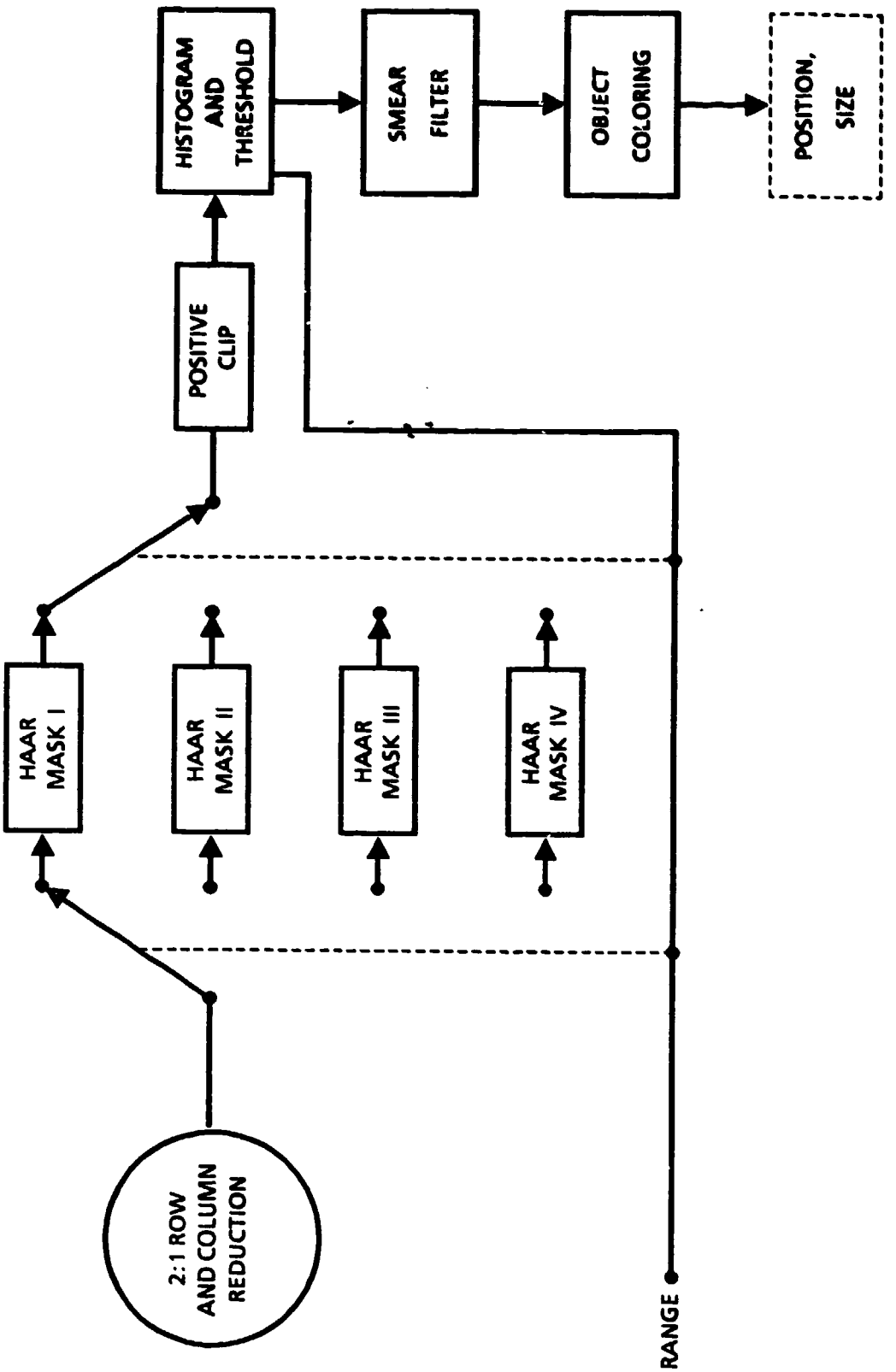


FIGURE E-1. HAAR DETECTOR

FIGURE E-2. ATC HAAR MASK IMPLEMENTATION

range to center field of view (CFOV). The following section outlines the sequence of algorithms used to post-process the filter output.

Histogram and Thresholding. In order to reduce the data bandwidth and extract meaningful information from the Haar transformed image, a decision must be made. This decision takes the form of a thresholding operation, yielding a binary resultant image. The threshold is computed as a percentile of the Haar image histogram, see Figure E-3. Determination of the threshold level is accomplished by down counting the histogram memory address and simultaneously comparing the summation of the stored frequency count to the desired percentile. The step–by–step outline for threshold generation follows:

(1) Start at the right of the histogram (highest pixel grey levels) and proceed to the left (lowest pixel grey levels). Count the number of pixels within bin and add this to a running sum.

(2) If the running sum divided by the total number of pixels in the frame equals or exceeds the given percentile then stop.

(3) Use the grey level associated with the current bin as the threshold.

The percentile is predetermined and fixed for each range bin. It was designed to yield a constant number of exceedences. The percentile is determined a priori as a function of the expected number of targets and the nominally expected target size. Typical percentile values are quite small and are usually specified to the nearest 0.01% (or, in fractional form, 1/10,000). For the ATC program, values of 16/10,000, 20/10,000, 30/10,000 and 30/10,000 were used for ranges 7km, 5km, 3 km, and 1.5km, respectively.

Smear Filter. The binary image generated by the thresholding operation contains points that have responded strongly to the Haar mask, indicating a relatively high degree of "spot" feature activity. An unfortunate artifact of the Haar mask convolution is the creation of sidelobes that, after thresholding, become separated from the main lobe. This can occur when there is a mismatch between the Haar mask size and target size, or if two or more targets are closely spaced. This results in a single target giving rise to multiple responses at different image locations. To help

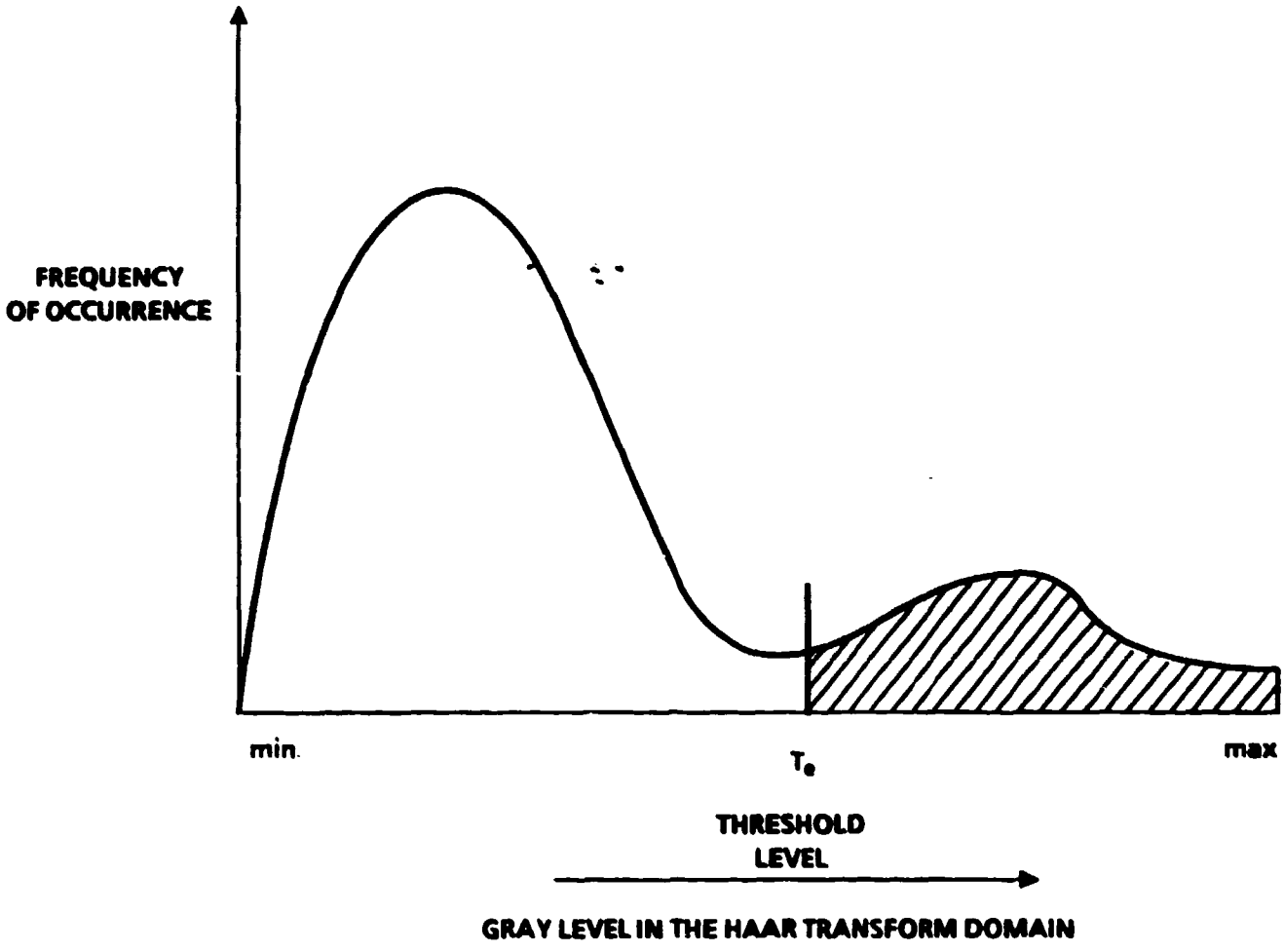


Figure E-3. TYPICAL HISTOGRAM OF HAAR PIXEL OUTPUT

ensure that disparate responses from a single target will be merged together, a 3 x 3 smearing operator is employed on the binary image (corresponding to a 6x6 operator on the original image). This simple blurring filter can be effected as an OR operation or as a max filter (equivalent operations when operating on a binary image).

Coloring. After blurring the binary image, a coloring operation is performed. The functions of the coloring process are to locate the centroid coordinates and to estimate the size of each candidate target. Object coloring is performed across the entire binary image plane based upon the 8 connectivity criterion. The coloring mask is shown in Figure E-4 along with a flow chart indicating mask rules and priorities. Figure E-5 contains a representative example of the coloring operation when applied to a binary image. The mask is convolved across the image, starting at the upper left-hand corner, scans across until the first line is complete, then retraces and scans the next line until the last line is completed. When the coloring mask is positioned over the pixel containing the first dot (binary 1), that pixel will be identified as belonging to cluster number 1. As the mask is slid across each line, the clusters will be numbered sequentially until mask element w, a, b, or c is positioned on a previously numbered cluster. If this occurs, a priority logic is invoked and the pixel is assigned the same cluster number as its highest priority neighbor. The example in Figure E-5 shows two pixels where conflicts occur (highlighted pixels) but are resolved by the priority logic.

After the convolutional process is completed and the conflicts resolved, the sizes and locations of the clusters are computed. This is accomplished by finding the maximum and minimum X and Y excursions of each cluster and performing a simple centroiding operation, respectively. The extraction of the position and size data concludes the postprocessing operation and marks the end of the detection (region of interest) process.

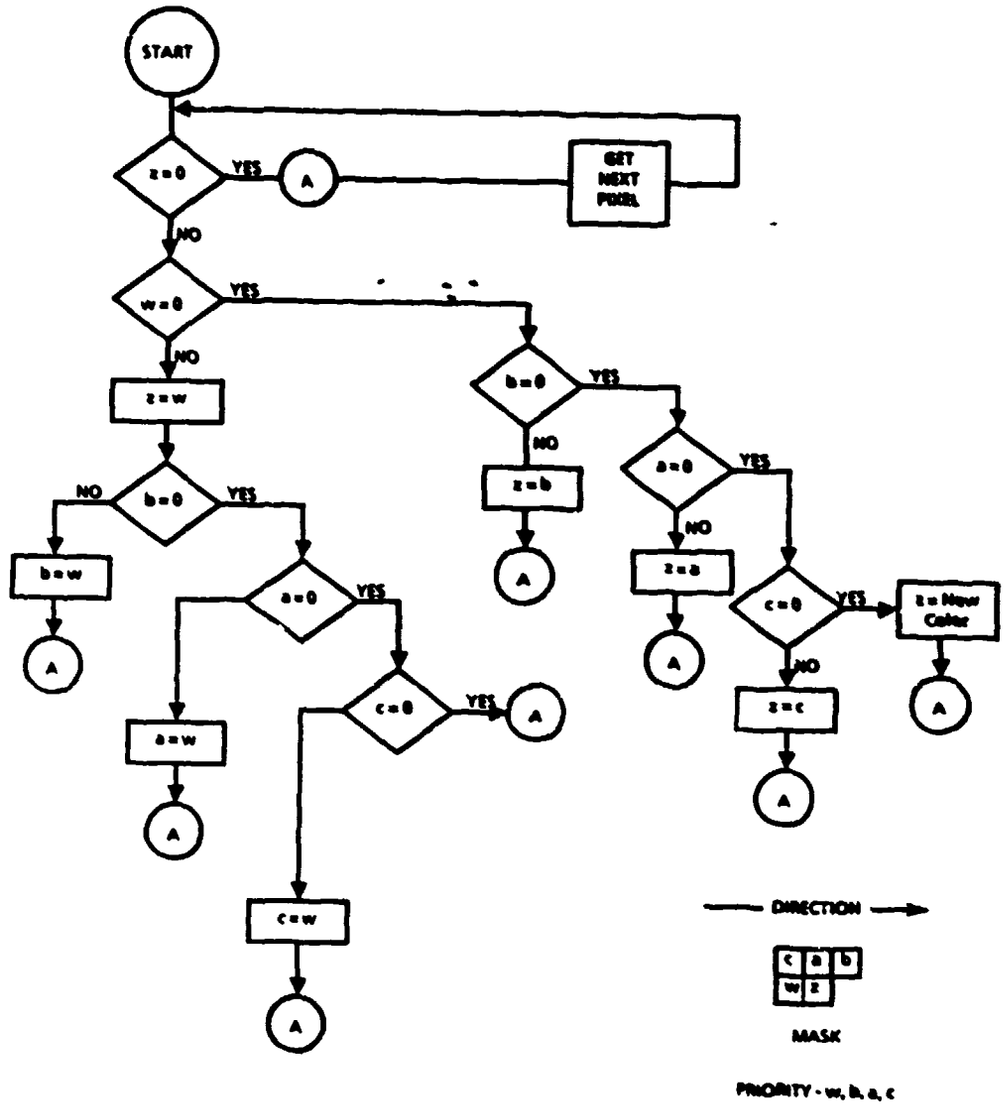
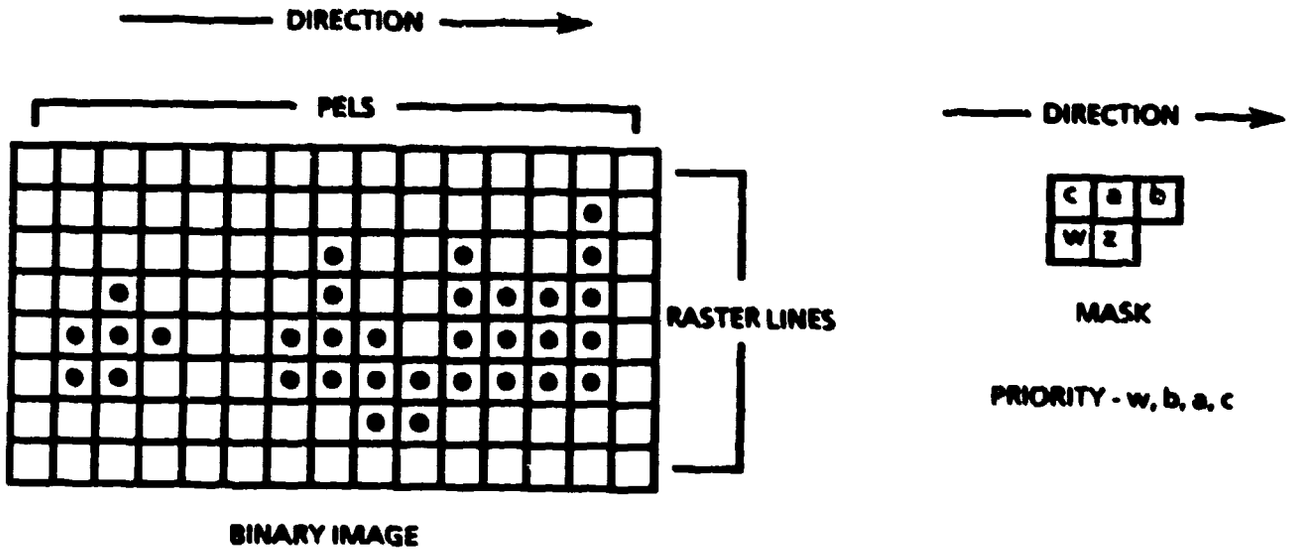
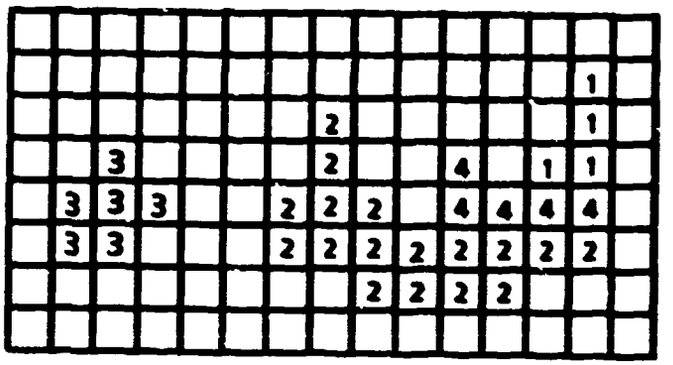


Figure E-4. COLORING FLOW DIAGRAM





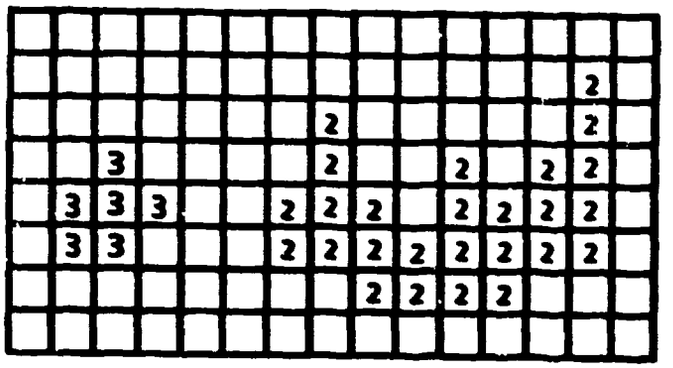


Figure E-5. COLORING LOGIC REQUIREMENTS

APPENDIX F

HAAR Figure of Merit

This appendix is a mathematical description of the Haar Figure of Merit (HFOM) as well as a brief overview of its functions and uses. The Automatic Target Cuer (ATC) simulation now uses the HFOM to quantify target detectability over the Haar transformed image. This metric can be used to attach a measure of detection confidence to each candidate target or as a discrimination variable.

The functional basis of the HFOM is the cross correlation operator. This type of operation is used extensively in image processing for tracking, registration, and template matching. The HFOM uses a HAAR mask as the reference template. The resulting cross correlation coefficient is now a normalized version of the Haar matched filter output. This normalization process enables the HFOM response to be gain and bias invariant. Its values range between plus and minus one and relate how well the Haar reference template matches various areas of the image. A value of +1 indicates perfect correlation, implying the highest possible detection confidence. A value of 0 (or less) indicates no (or negative) correlation and is a firm basis for rejection as a possible detection. The following compares the conventional Haar filtering process and the Haar Figure of Merit.

The Haar filter output is defined to be:

$$H(i,j) = \sum_{m} \sum_{n} [Mask(m,n) \bullet Image(i+m, j+n)] \quad , \tag{F-1}$$

where

i, j	=	Row and column coordinates of the center of the mask in image coordinates
n,m	=	Row and column coordinates of the Haar mask in mask coordinates

Mask = **Haar mask**

Image = **Input image**

The Haar Figure of Merit is defined to be:

$$HFOM(i,j) = \sum_m \sum_n \frac{[Mask(m,n) \bullet (Image(i+m, j+n) - \mu ij)]}{(\sigma_{ij}^2 \sigma_m^2)^{\frac{1}{2}}} \quad \text{(F-2)}$$

where

$$\mu_{ij} = \frac{1}{mn} \sum_m \sum_n Image(i+m, j+n) \quad \text{(F-3)}$$

$$\sigma_{ij}^2 = \sum_m \sum_n [Image(i+m, j+n) - \mu_{ij}]^2 \quad \text{(F-4)}$$

$$\sigma_m^2 = CONSTANT = \sum_m \sum_n [Haar(m,n)]^2 \quad \text{(F-5)}$$

APPENDIX G

Relationship Between Kendall's τ and the Sample Correlation Coefficient

The following equation is easily recognized as the formula for r_s, the sample correlation coefficient,

$$r_s = \frac{\sum (x_i - x)(y_i - y)}{\left(\sum (x_i - x)^2 \sum (y_i - y)^2\right)^{\frac{1}{2}}} \qquad \text{(G-1)}$$

The formula for r_s can be considered as a special case of the more general correlation measure, Γ, where Γ is defined*

$$\Gamma = \frac{\sum a_{ij} b_{ij}}{\left(\sum a_{ij}^2 \sum b_{ij}^2\right)^{\frac{1}{2}}} \qquad \text{(G-2)}$$

where a_{ij} and b_{ij} are functions of (x_i,x_j) and (y_i,y_j) that replace and augment the original x_i and y_i data points. (r_s is derived from Γ by letting $a_{ij} = 0$ if $i \neq j$, and x_i-x if $i = j$, and similarly for the b_{ij} and y_i.) It can be shown that Γ always lies in the interval [-1,1] no matter what the values of a_{ij} and b_{ij} (the Schwarz Inequality), with $\Gamma = 1$ when the $\underline{a}$ and $\underline{b}$ vectors are identical and $\Gamma = -1$ when the $\underline{a}$ vector is -1 times the $\underline{b}$ vector. The purpose in generalizing the correlation coefficient is to consider alternative measures of correlation and evaluate the trade-offs between different formulations to choose the method which best suits the application.

One very useful generalization is to let a_{ij} and b_{ij} be functions of the relative rank values of x's and y's, i.e., $a_{ij} = 1$ if $x_i < x_j$, let $a_{ij} = 0$ if $x_i = x_j$, and let $a_{ij} = -1$ if $x_i > x_j$. If we first rearrange the y's to be strictly increasing, that is

$$y_i < y_j, \quad for j > i, \qquad \text{(G-3)}$$

and let the associated x's take on the resultant ordering, then if we limit the calculations to $j > i$ we have

$$\Gamma = \frac{\Sigma a_{ij}\, b_{ij}}{\left(\Sigma a_{ij}^2 \Sigma b_{ij}^2\right)^{\frac{1}{2}}} = \frac{\Sigma a_{ij} \cdot 1}{\Sigma a_{ij}^2 \Sigma 1^2}. \qquad \text{(G-4)}$$

---

*Kendall, M. G., Rank Correlation Methods, Charles Griffin Co. Ltd., London, 1962.

$$So \qquad \Gamma = \frac{\Sigma a_{ij}}{\left[\left(\Sigma a_{ij}^2\right) \cdot \frac{N(N-1)}{2}\right]^{\frac{1}{2}}}, \tag{G-5}$$

$$or \qquad \Gamma = \frac{\sum_{X_i < X_j} a_{ij} + \sum_{X_i = X_j} a_{ij} + \sum_{X > X_j} a_{ij}}{\left[\sum_{X_i \neq X_j} a_{ij}^2 + \sum_{X_i = X_j} a_{ij}^2\right]^{\frac{1}{2}} \left[\frac{N(N-1)}{2}\right]^{\frac{1}{2}}} \tag{G-6}$$

In terms of Kendall's parameters, this becomes

$$\Gamma = \frac{T + 0 - I}{\left[\frac{N(N-1)}{2}\right]^{\frac{1}{2}} \left[\frac{N(N-1)}{2}\right]^{\frac{1}{2}}}, \tag{G-7}$$

$$or \qquad \Gamma = \frac{T - I}{\frac{N(N-1)}{2}} = \tau. \tag{G-8}$$

Both τ and r_s are measures of correlatedness. (In fact, it can be shown that, for the case where the x's and y's have Gaussian distributions, $E(\tau) = (2/\Pi) \sin^{-1} r_s$.) However, tests for significant values of r_s depend on the assumption that the x's and y's are Gaussianly distributed, whereas tests based on τ do not depend on the underlying distributions of the x's and y's for evaluation of significance.

APPENDIX H

Numerical Example of Kendall's τ and Associated Parameters

The following is a numerical example of how Kendall's τ and associated parameters are determined. The data used are the same as those used to produce the results in the last row of Table 10 in the main body of this report. Table H-1 lists the data. Four quantities are calculated for these data: Kendall's τ; the upper and lower bounds, τ_u and τ_l, of a 95% confidence interval containing τ; and α, the probability under the null hypothesis of obtaining a value as large as τ. The International Mathematics and Statistics Library (IMSL) subroutine NMKN was used in computing τ and α.

Calculation of τ

In Section 5.1.6 of the main body of this report Kendall's τ was defined as the ratio of (T-I) to n (n-1)/2, where T is the number of data point pairs exhibiting positive monotonicity, I is the number of data point pairs exhibiting negative monotonicity, and n is the number of data points. For this data set, T-I = 1292 and n = 52, so τ = 0.974359.

Calculation of Confidence Interval

The formula for the endpoints of a 95% confidence interval for τ is (Kendall, 1962)

$$\tau_l = \frac{\tau - z\sqrt{2/n}\,\sqrt{1 + z^2/n - \tau^2}}{1 + 2z^2/n} \qquad \text{(H-1)}$$

$$\tau_u = \frac{\tau + z\sqrt{2/n}\,\sqrt{1 + z^2/n - \tau^2}}{1 + 2z^2/n} \qquad \text{(H-2)}$$

where z is the value of a standard normal variable (with zero mean and unit variance) such that the area to the right of z under the standard normal curve is 0.025. For this data set, and for a 95% confidence interval, $\tau_l = .700$, $\tau_u = .998$

Calculation of α

The distribution of $(T\text{-}I)/\sqrt{Var(T\text{-}I)}$, under the hypothesis data that the data exhibit no correlation, is approximately normally distributed with zero mean and unit variance for n greater than 7 (Kendall, 1962). The formula for Var (T-I), when there are no ties in the data (as is the case with the data in Table 1), is

$$Var(T-I) = n(n-1)(2n+5)/18. \text{ (Kendall, 1962)}$$

(Kendall also provides the formula for the case of ties). For the given data set, n = 52, Var(T–I) = 16,060, so under the hypothesis of no correlation,

$$Pr\left(\frac{T-I}{\sqrt{Var(T-I)}} > \frac{1292}{\sqrt{16,060}}\right) = 1 \bullet 10^{-24}$$

TABLE H-1. Integral Measure Data for TIR on Small Target Image Set

DATA POINT	AVERAGE TIR	HAAR FOM
1	6.159593	0.5614600
2	5.469493	0.5391429
3	2.219200	0.4323732
4	12.35137	0.6026000
5	3.243192	0.4773111
6	2.222797	0.4323350
7	2.374016	0.4340539
8	4.187746	0.5272800
9	2.850844	0.4519967

DATA POINT	AVERAGE TIR	HAAR FOM
10	3.439575	0.4838438
11	3.070657	0.4630913
12	4.057651	0.5072364
13	2.888486	0.4571069
14	2.960361	0.4556424
15	3.103345	0.4664000
16	7.193048	0.5782000
17	3.054875	0.4597251
18	3.848121	0.4903857
19	10.75095	0.6036000
20	3.143448	0.4735286
21	5.776412	0.5436667
22	3.225395	0.4710947
23	3.887485	0.4954538
24	2.160655	0.4295143
25	7.139445	0.5652000
26	2.125995	0.4270907
27	2.493509	0.4388676
28	0.9968202	0.4125404
29	2.686939	0.4494876
30	2.540206	0.4467971
31	1.276870	0.4161980
32	1.754863	0.4214500
33	1.321635	0.4203938
34	3.289654	0.4718824
35	2.537872	0.4445744
36	5.455533	0.5450125
37	3.689721	0.4859934
38	2.958680	0.4556778
39	2.533487	0.4409723
40	4.497394	0.5443000
41	2.615864	0.4480183
42	2.030863	0.4246614
43	1.308730	0.4187021

DATA POINT	AVERAGE TIR	HAAR FOM
44	1.030863	0.4142040
45	2.399988	0.4353869
46	1.766195	0.4242601
47	1.597554	0.4227596
48	3.016114	0.4580561
49	3.991664	0.5002417
50	2.831542	0.4504904
51	3.145221	0.4709350
52	2.918194	0.4560429

Kendall, M. G., Rank Correlation Methods, Charles Griffin Co. Ltd., London, 1962.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: 321963	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): Image Metrics	5. TYPE OF REPORT & PERIOD COVERED: Scientific & Technical Interim 10/85 - 4/86
6. PERFORMING ORG. REPORT NUMBER: 86Y061	
7. AUTHOR(s): A.B. Lucero, G.B. Silverman, J.W. Bair, R.R. Ramroth	8. CONTRACT OR GRANT NUMBER(s): DAAL02-85-C-0165
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Northrop Corporation, EMD 500 East Orangethorpe Avenue Anaheim, CA 92801	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: CLIN 0001 Data Item A004
11. CONTROLLING OFFICE NAME AND ADDRESS: DCASMA - Santa Ana P.O. Box C - 12700 Santa Ana, CA 92712	12. REPORT DATE: 1 May 1986
	13. NUMBER OF PAGES: 88
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): Same as Block 11	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

(3) W26P70
(2) DARPA/TTO
(1) DARPA/TIO
(30) AMSEL-NV-AC

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

Not different from report distribution

18. SUPPLEMENTARY NOTES

None

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Image Metrics
Image Quality
Signal Quality
Forward Looking Infrared
Automatic Target Recognizer Evaluation
Detection
Ground Targets
Ground Clutter
Target to Interference Ratio

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

In order to respond to a DoD need for a means to objectively and quantitatively measure image signal quality for evaluation of automatic target recognizers (ATR's) with respect to target detectability and clutter rejectability, this study was initiated. It builds upon prior work in this area. Target to interference ratio (TIR) and other metrics are investigated. A validation exercise is conducted in which image metrics are correlated against the Haar detector that is used in the Northrop/NVEOC Automatic Target Cuer (ATC). Excellent correlation is obtained for TIR and another image metric, for both targets and clutter objects.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)